TITLE OF THE INVENTION

A SIZE-VARIABLE STRAIN-SPECIFIC PROTECTIVE ANTIGEN FOR POTOMAC HORSE FEVER

This application claims the benefit of priority under 35 U.S.C. §119(e) to U.S. Provisional Application Serial No. 60/059,252, filed on September 18, 1997.

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates to isolated and purified antigen which is expressed by a wild-type *E. risticii* strain and is specific to the strain. The present invention also relates to nucleic acid constructs which encode the antigen, expression vectors, transformed host cells, and methods for producing the antigen.

Discussion of the Background

Potomac horse fever (PHF), also known as equine monocytic ehrlichiosis (EME), is an acute infectious disease of horses. PHF was initially recognized in 1979 in areas along the Potomac river in Maryland and Virginia. The causative agent was subsequently identified in 1984 as *Ehrlichia risticii*, an obligatory intracellular rickettsial organism. Since then, PHF cases have been reported in many states of the U.S. and some provinces of Canada. Serological evidence suggests the presence of *E. risticii* in parts of Europe and Australia. The main disease features of PHF are fever, leukopenia, depression, anorexia and diarrhea. Some affected horses may also develop colic or laminitis. The mortality is as high as 20-25%. Recently, abortions in pregnant mares contracting the disease have been documented. PHF occurs mostly in the summer months. Although most of the rickettsial pathogens are transmitted by arthropod vectors and the seasonality of PHF also suggests this, all attempts to

reveal the mode of transmission of *E. risticii* have been unsuccessful.

E. risticii infection is responsible for substantial economic loss to the equine industry. Currently, inactivated vaccines for PHF are commercially available from three different manufacturers. In endemic areas, vaccination of equine population against PHF is performed on a regular basis. Despite this, PHF is occurring in increasing numbers, including in vaccinated horses. In 1990, *E. risticii* was isolated from a horse suffering from severe PHF in spite of carrying a high titer of antibodies from multiple PHF vaccinations. On Western blot analysis, the antigenic profile of this newly isolated organism (90-12 strain) was considerably different from that of the original organism (25-D strain) isolated in 1984 during the initial outbreaks of the disease. In subsequent years, more isolates were obtained from vaccinated horses suffering from clinical PHF. These findings suggested the possible existence of strain variation in *E. risticii* and its probable role in vaccine failures in the field.

In the last few years, significant progress has been made toward understanding the pathogenesis and host immune response in *E. risticii* infection. Certain strains of mice have been identified to be good laboratory models of PHF. Various serological and DNA based tests have been developed to better facilitate diagnosis of the infection. Studies to identify the antigenic composition of the organisms and the major surface antigens involved in immune response were conducted. However, most of these studies have been performed with the original *E. risticii* isolates (isolated during 1984-85) from different laboratories. Except for one recent report on biological diversity in *E. risticii* isolates, no systematic comparison between different isolates has been made to identify the extent and importance of strain variation in this organism. Also, very little is known about the molecular biology of *E. risticii*. Hence, the present study has been undertaken to: i) understand the differences between the 25-D and 90-12 strains of *E. risticii*; ii) investigate the molecular basis of these differences; iii) identify protective antigen(s).

In addition to the main focus of problem solving *E. risticii* infections, there is an important scientific interest in these studies to gain more knowledge on ehrlichial organisms in general. Along with *E. risticii*, genus *Ehrlichia* of the family *Rickettsiaceae* contains some other recently identified organisms. New members of this genus include *E. chaffeensis* and *E. ewingii*, pathogens of human and dog, respectively. Recently identified human

granulocytic ehrlichiosis (HGE) has been demonstrated to be caused by an organism similar to or the same as *E. equi*, an equine pathogen. Also, *E. risticii* has been found to infect dogs and cats. Emergence of these ehrlichial diseases and changes in host specificity of ehrlichial organisms are quite intriguing. Information on the important proteins of *E. risticii* and the genes they are encoded by may provide us with necessary clues to understand the sophisticated intracellular survival strategies of ehrlichial organisms and the natural dynamics in their ecosystem that lead to changes in their life cycles.

SUMMARY OF THE INVENTION

The present invention is based on the discovery that strains of *Ehrlichia risticii* express surface antigens that are specific to the strain. These surface-expressed proteins are termed strain-specific antigens (SSAs). These antigens have now been isolated and purified from the respective strains. The SSAs of the present invention may be used to detect *Ehrlichia risticii* strains and to generate a protective immune response against *E. risticii* strains, leading to the development of more effective vaccines against PHF.

BRIEF DESCRIPTION OF THE DRAWINGS

A more complete appreciation of the invention and many of the attendant advantages thereof will be readily obtained as the same becomes better understood by reference to the following detailed description when considered in connection with the accompanying drawings.

Figure 1: Primers sequences ECP 1 and ECP 2 for amplifying any SSA gene (SEQ ID NO: 1 and 2).

Figure 2: Nucleotide sequence of 85kD gene with flanking regions and deduced amino acid sequence of the strain-specific antigen from *E. risticii* 90-12 strain (SEQ ID NO: 3 and 4). Putative -10, -35, RBS regions are underlined and putative starts of transcription is denoted (+1). The dyad symmetry, and the adjacent thymine-rich regions are underlined.

Figure 3: Nucleotide sequence of 50kD gene with flanking regions and deduced amino acid sequence of the strain-specific antigen from *E. risticii* 25-D strain (SEQ ID NO: 5

and 6). Putative -10, -35, RBS regions are underlined and putative starts of transcription is denoted (+1). The dyad symmetry, and the adjacent thymine-rich regions are underlined.

Figure 4: Nucleotide sequence of ATCC-50kD gene with flanking regions and deduced amino acid seqüence (SEQ ID NO: 7 and 8). Putative -10, -35, RBS regions are underlined and putative starts of transcription is denoted (+1). The dyad symmetry, and the adjacent thymine-rich regions are underlined.

Figure 5: Analysis of deduced amino acid sequences of SSA homologues from two antigenic variants of *E. risticii* (SEQ ID NO: 4 and 6). There were a total of eight identical domains present in both 50kD and 85kD antigens. The number at the top show each identical domain. There were significantly high homology present in the corresponding domain of the same number. Minor amino acid changes in each domain in 85kD identified after compared with 50kD and marked by a black triangle head.

Figure 6: Pre-challenge serum antibody titers of mice from different groups of experiment 1. All the antigens used for immunization of mice were from the 90-12 strain. Antibody titers were determined by performing indirect immunofluorescent assay (IFA). MM cells infected with the 90-12 strain were used in the IFA.

Figure 7: Pre-challenge serum antibody titers of mice from different groups of experiment 2. Antibody titers were determined by performing IFA using MM cells infected with the 90-12 strain.

Figure 8: Post-challenge clinical signs of mice from different groups of experiment 2. Clinical signs were scored on a scale of 0 to 5, with 5 representing the most severe symptoms.

Figure 9: DAF patterns (size of the amplified DNA in ethidium bromide agarose gel electrophoresis) of field strains of *E. risticii*. Group 1 (1.88kb): Isolates 94-2, 94-3, 94-24, 90-30 and 25-D strain in lanes 3, 4, 5, 7 and 2. Group 2 (1.86kb): Isolate 94-27 in lane 6. Group 3 (1.80kb): Isolate 94-28 in lane 13. Group 4 (1.75kb): Isolates 94-8, 94-31, 94-37, 94-49 and 94-50 in lanes 12, 13, 15, 16 and 17 which is similar for Illinois/ATCC strain (1.75kb). Group 5 (1.56kb): Isolate 64--29 and 90-12 strain in lanes 11 and 8. Group 6 (1.45kb): Isolates 94-22 and 94-25 in lanes 9 and 10. The DNA from uninfected mouse

macrophage cells were used as a control in PCR amplification (lane 18). No visible band in lane 18 indicates the specificity of the primers. Molecular weight markers in lanes 1 and 19.

DETAILED DESCRIPTION OF THE INVENTION

As used herein, the term "isolated and purified" refers to an antigen that has been separated and isolated from the *E. risticii* strain expressing the same. Preferably, the inventive SSA is separated from other proteins derived from *E. risticii*, especially other antigenic proteins.

The SSAs of the present invention may be obtained via PCR amplification from the genomic DNA of a wild-type *E. risticii* strain using well-known molecular biology techniques. Such techniques are well-known to those skilled in this art. The oligonucleotide primers for isolating a desired SSA gene may be prepared based on the specific nucleotide sequences disclosed herein. Specific examples of suitable primers are shown in Figure 1 (SEQ ID NO: 1 and 2). For a discussion of PCR amplification, see *Current Protocols in Molecular Biology*, F. M. Ausubel et al, Eds., Volumes 1-3, John Wiley and Sons, 1998, incorporated herein by reference.

The SSA may vary widely in both overall size and amino acid composition. The SSA may have a molecular weight of about 40 to about 90 kDa, inclusive of all specific values and subranges therebetween. In specific embodiments of the present invention, the SSA has a molecular weight of about 50 kDa or 85 kDa. Examples of specific amino acid sequences of the inventive SSAs are shown in Figures 2-4 (SEQ ID NO: 3, 5 and 7).

The present invention also provides isolated and purified nucleic acids (e.g., recombinant DNAs) which encode the SSAs. Specific examples of nucleotide sequences encoding the SSA of the present invention are shown in Figures 2-4 (SEQ ID NO: 4, 6 and 8). All nucleotide sequences encoding a particular SSA are included in the scope of the present invention. Selecting a nucleic acid encoding a particular amino acid sequence may be readily accomplished using the well-established genetic code relating the nucleic acid sequence of a codon sequence to the amino acid sequence encoded thereby. The genetic code is provided

by R. H. Abeles et al, Biochemistry, Jones and Bartlett, 1992, p. 269, incorporated herein by reference in its entirety.

All percentage identities for the amino acid and DNA sequences noted above can be determined using a variety of algorithms known in the art. An example of a useful algorithm in this regard is the algorithm of Needleman and Wunsch, which is used in the "Gap" program by the Genetics Computer Group. This program finds alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps. Another useful algorithm is the algorithm of Smith and Waterman, which is used in the "BestFit" program by Genetics Computer Group. This program creates an optimal alignment of the best segment of similarity between two sequences. Optimal alignments are found by inserting gaps to maximize the number of matches using the local homology algorithm of Smith and Waterman. It is preferred to use the algorithm of Needleman and Wunsch to compare the amino acid and DNA percentage identity in the present case to another amino acid or DNA sequence.

The nucleic acid encoding the SSA may be incorporated into a vector suitable for directing the expression of the SSA in a suitable host (i.e., recombinant expression). Such expression vectors may have all of the customary transcriptional control elements which enable the SSA to be expressed in a host transformed with the vector. For a detailed discussion of expression vectors and related cloning technology, see *Current Protocols in Molecular Biology*, supra.

Suitable host cells include bacteria customarily used in the overproduction of recombinant protein sequences, e.g., *E. coli*. Mammalian cells may also be used as host cells if desired.

The inventive SSA may be produced by culturing a host cell transformed with an expression vector carrying the nucleic acid encoding the antigen in a suitable culture medium. The antigen is then isolated from the culture medium according to well-known procedures.

The isolated and purified SSA may be formulated into an immunogenic pharmaceutical composition by incorporating an effective amount of the antigen into a pharmaceutically acceptable carrier. Suitable carriers include, for example, aqueous solutions

containing the customary components for administration to host, e.g., buffers, salts, adjuvants, etc. Upon administration of the composition to a host, the antigen induces a protective immune response against the *E. risticii* strain from which the antigen was derived. Preferably, such an immunogenic composition is a vaccine against the wild-type *E. risticii* strain from which the antigen was derived. Of course, in a preferred embodiment, the antigen also produces an immune response against other strains besides the wild-type strain from which the antigen is derived. In other words, a SSA from one strain may contain one or more epitopes which are shared with the SSA of other strains. A suitable host for the inventive immunogenic composition is, for example, a horse. The host may be any other animal that is susceptible to infection by *E. risticii* (e.g., cats, dogs and humans). Formulating immunogenic pharmaceutical composition, administering the composition to a host, and determining the level of induced immune response are readily accomplished using techniques well-known to those skilled in this art.

Having generally described this invention, a further understanding can be obtained by reference to certain specific examples which are provided herein for purposes of illustration only and are not intended to be limiting unless otherwise specified.

EXAMPLES

Example 1

Isolation of Strain Specific Surface Antigen Gene of *Ehrlichia risticii*

Ehrlichia risticii strains and DNA preparation:

Two different strains of *E. risticii* were used for this study. The original *E. risticii* strain (25D) was isolated in 1984, during the initial outbreaks of PHF near the Potomac River bank in Maryland and Virginia. Recently the inventors isolated a new strain of *E. risticii* (90-12) from a vaccinated horse suffering from clinical PHF. These two strains of the organism were grown separately in human histiocyte cells and purification was accomplished over linear Renografin (Squibb, NJ) density gradient centrifugation.

Propagation of *Ehrlichia risticii* in cell culture:

E. risticii strains were propagated a human histiocyte (HH) cell line (American Type culture Collection # U937). These cells were grown in RPMI 1640 medium (Flow Laboratories, McLean, VA), supplemented with 4 mM L-glutamine (M.A. Bioproducts, Walkersville, MD), and 15% fetal calf serum (Gibco Laboratories, Grand Island, NY). Approximately 20×10^6 cells in the logarithmic phase of growth were centrifuged at 500×g for 10 minutes and the cell pellet was resuspended in 20 ml of *E. risticii* infected HH cell culture. This infected cell mixture was dispensed into a 150 cm^2 tissue culture flask and incubated at 37°C in a humidified chamber in the presence of 5% CO_2 for one hour. Seventy ml of the growth medium was then added and the culture was incubated further.

The infected cell cultures were examined for infection by acridine orange staining according to a standard procedure. For this, about one ml of the cell suspension was centrifuged at 500×g for five minutes. The cell pellet, resuspended in about 50 µl of the supernatant, was applied onto a glass slide and allowed to air dry. The cells were fixed with absolute methanol for 10 minutes, stained with acridine orange stain for three minutes and examined with an ultraviolet microscope. The efficiency of infection of *E. risticii* was determined by considering the number and the intensity of orange specks of *E. risticii* inside the pale green stained cytoplasm of the HH cells.

The infected cultures were harvested on day 4-6 postinfection, depending upon the observed levels of infection. Infected HH cultures were centrifuged at 17,000×g for 20 minutes in a Sorvall refrigerated centrifuge (Sorvall, Norwalk, CT) and the cell pellet was resuspended in sodium-potassium-glutamine buffer to a final concentration of 50× and stored at -70°C.

Purification of *Ehrlichia risticii:*

E. risticii organisms were purified by centrifugation over a linear Renografin gradient according to known procedures described. Ten ml of 50× concentrate of infected HH culture were diluted with 20 ml of Tris buffer (10 mM Tris, pH 7.4). All the buffers contained 1.0 mM phenylmethylsulfonylfluoride (PMSF) and 1.0 mM iodoacetamide as proteinase

inhibitors. The cell suspension was homogenized for three cycles in a Omni mixer (Dupont Co., Wilmington, DE) at maximum setting for 30 seconds. Each cycle was followed with 30 seconds of cooling on ice. The homogenate was clarified at 2,000×g for 10 minutes to sediment the nuclear material and unbroken cells. The supernatant was centrifuged at 17,000×g for 20 minutes. The resulting pellet was resuspended in 2.0 ml of Tris buffer and the suspension was forced through 18 and 23 gauge needles to obtain homogeneity. The volume was brought up to 12 ml with Tris buffer containing 10 mM $MgSO_4$. Two µl (20 µg) of Dnase I (Life Technologies, Inc., Gaithersburg, MD) were added to the suspension and incubation carried out at 37°C for about 10 minutes to digest the liberated host nuclear material. Two ml of the suspension were then layered on about 34 ml of a 20 to 45% linear density gradient of Renografin (Squibb Chemical Co., New Brunswick, NJ) in TEN buffer (50 mM Tris, pH8.0, 25mM EDTA, 0.9% NcCl). The gradients containing ehrlichiae were centrifuged at 83,000×g for one hour at 4°C. The ehrlichiae were observed to band at a density of about 1.182 gm/ml and could be visualized well with a pointed light source. The cellular debris formed a compact band at the top of the gradient. The ehrlichial bands were collected and diluted with 10 volumes of Tris buffer and centrifuged at 17,000×g for 20 minutes to remove the Renografin. The purified *E. risticii* pellet then was resuspended in TEN buffer to a final concentration of 500× for the DNA experiments or to a final concentration of 200× in 10 mM tris buffer, pH7.4, for all the analyses.

Extraction of *Ehrlichia risticii* DNA:

Purified *E. risticii* suspension (500×) in TEN buffer was treated with lysozyme (Sigma Chemical Co., St. Louis, MO) at a final concentration of 2.0 mg/ml and incubated in a 37°C water bath for 30 minutes. To this digest, SDS was added to a final concentration of 0.5% and the lysate was kept in a 65°C water bath for an additional 30 minutes. This lysate was then treated with proteinase K (Bethesda Research Laboratories., Gaithersburg, MD) at a final concentration of 400 mg/ml and was incubated in a 56°C water bath overnight. Two phenol extractions were done with equal volumes of water saturated phenol and by shaking on a Orbitron rotator (VWR Scientific, Brisbane, CA) for 30 minutes each. Three chloroform

(chloroform:isoamyl alcohol; 24:1) extractions were done and the DNA was precipitated by the addition of sodium acetate (pH 5.2) to a final concentration of 0.3M. Two volumes of absolute ethanol was added and incubation was carried out at -20°C for two hours. The precipitate was pelleted by centrifugation at 12,000 rpm in a table top microcentrifuge for 15 minutes at 4°C. The DNA precipitate was washed once with 70% ethanol and then with absolute ethanol each followed by centrifugation at 4°C. The DNA pellet was allowed to dry in a vacuum for five to ten minutes and then it was dissolved in TE buffer (10 mM Tris, pH 8.0, 1.0 mM EDTA) to a concentration of 1.0 µg DNA/µl and stored at -20°C for future use.

Polyacrylamide gel electrophoresis and Western immunoblotting:

Discontinuous SDS PAGE analyses were carried out over 10 and 12% polyacrylamide gels according to the method of Laemmli (120). Gels were cast on a vertical slab gel electrophoresis system (Model SE 600, Hoeffer scientific Instrument, San Francisco, CA). For a 10% gel, 10 ml of acrylamide solution containing 30% acrylamide and 2.7% N,N′ -methylene bisacrylamide were mixed with 7.5 ml of 1.5 M Tris buffer, pH 8.8, 150 µl of 20% SDS, and 10.5 ml of distilled water and degassed for 15 minutes under vacuum. Polymerization was initiated by the addition of 150 µl of 10% ammonium persulfate and 10 µl TEMED (N,N,N′ ,N′ -tetramethylethylene- diamine), and then the solution was poured immediately into the gel apparatus. About 1.0 ml of water saturated butanol was layered on top and the gel was allowed to polymerize for about 30 minutes. The stacking gel contained 1.33 ml of acrylamide solution, 2.5 ml of 0.5 M Tris buffer, pH 6.8, 100 µl of 10% SDS, 6.1 ml of water, 50 µl of 10% ammonium persulfate and 5.0 µl of Temed. The samples were dissolved in sample buffer (62.5 mM Tris, pH6.8, 2% SDS, 5% 2-mercaptoethanol, 10% glycerol) and heated to 100°C for 10 minutes, and loaded onto the gel. The gels were electrophoresed for 1111 volthours at a constant current with an automated power supply (model 3000 xi, Bio-Rad laboratories, Richmond, CA), while the apparatus was kept cooled to 4°C using a thermostatic circulator (LKB Instruments, Bromma, Sweden). The gels were stained with 0.05% Coomassie blue in 40% methanol, 10% acetic acid, or processed for Western

blotting.

The Western immunoblotting was conducted according to the method of Towbin *et al.* using a transfer apparatus (Hoffer). The Western blotting sandwich contained 3.0 mm Whatman filter paper (Whatman Limited, England), nitrocellulose membrane (NCM; Bio-Rad), polyacrylamide gel, and 3.0 mm filter paper in that order. The sandwich was assembled in a tray containing blotting buffer such that no air bubbles were trapped between the sandwich layers. The transfer was performed at 100 volts for four to six hours with a transfer power supply (Hoeffer). The temperature was maintained at 4°C during the transfer using a thermostatic circulator.

After the transfer, the NCM were cut, using the pre-stained molecular weight marker (Bio-Rad) lane as a guide, and the unbound sites were blocked by incubation in a two percent casein solution for three hours at 4°C. The antibodies were diluted in two percent casein solution and incubated with the membranes in a 150 mm diameter petri dishes or in hybridization bags (BRL) for three hours at room temperature, or for overnight at 4°C. The membranes were washed twice in Tris saline (10 mM Tris, pH 7.4, 150 mM NaCl), with 0.05% Triton X-100, and once in Tris saline for 15 minutes each. The membranes were then incubated with the appropriate alkaline phosphatase labeled antibodyies (Kirkegaard and perry laboratories, Inc.,) diluted to 2.0 μg/ml in casein solution, for one hour at room temperature. The membranes were washed as described earlier, followed by a final wash with 0.9% NaCl. Color development was accomplished with Fast Red TR salt and napthol AS MX phosphate substrates for about 10 minutes, and the reaction was stopped by washing the membrane in distilled water. The diluted sera and enzyme-labeled antibody solutions were stored at -70°C for reuse.

Cloning of *Ehrlichia risticii* genomes of original (25D) and variant (90-12) strains:

Fragments of the genomic DNA of *E. risticii* (25D strain) were molecularly cloned in λ-gt11 vectors and a recombinant expressing a complete 50kD protein antigen gene was identified. Additional cloning of *E. risticii* (90-12 strain) was performed with similar procedures in λ-ZAP (Stratagene, LA Jolla, CA) as described below

Construction of variant *Ehrlichia risticii* recombinants:

Restriction enzymes were obtained from Bethesda Research Laboratories (Gaithersburg, MD), Promega Corporation (Madison, WI) and New England Biolabs (Beverly, MA). T4 DNA ligase, λ packing mix, λ-ZAP II, pBluescript phagemids, and *E. coli strain* XL1-Blue [*recA1 endA1 gyrA96 thi hsdR17* (r_k^- m_k^-) *supE44 relA1* λ⁻ *(lac) {F' proAB lac,*^{q} *Z M15,* Tn10*(tet*^{r}*)}*], were obtained from Stratagene (La Jolla, CA).

Restriction enzyme digestion of *Ehrlichia risticii* DNA:

Variant *E. risticii* genomic DNA was restriction digested by using Sau3AI (New England Biolabs) site-specific endonuclease in the following manner: Six µl of DNA sample containing 1.0 µg/µl of DNA were mixed with 36 µl of distilled water and 5.0 µl of 1× Sau3A I digestion buffer [100 mM NaCl, 10 mM Tris-HCL, 10 mM $MgCl_2$, (pH 7.3)], supplemented with 0.5 µl (100 µg/ml) bovine serum albumin. The contents of the tube were gently mixed in an eppendorf centrifuge at 10, 000 rpm for five seconds. Finally, 2.5 µl of enzyme (10 units/µl) were added and the mixture was again centrifuged at 10,000 rpm for five seconds in an Eppendorf centrifuge, and was kept at 37°C in a water bath for one hour. The reaction was stopped by the addition of EDTA to a final concentration of 25 mM. A small aliquot was electrophoresed over 1% agarose gel to monitor the digestion. One hundred µl of TE buffer were added to the mixture and the DNA was extracted once with phenol and subsequently washed three times with chloroform:isoamyl alcohol at the ratio of 24:1. The restriction digested DNA was precipitated with ethanol as described above.

Synthesis and ligation of adapters to *Ehrlichia risticii* DNA fragments:

Three different types, (1, 2, and 3) of EcoR I-BamH I conversion adapters were prepared by the annealing of six different kinds of synthetic oligonucleotides, and each of these adapters was ligated separately to the Sau3A I cohesive ends of the variant *E. risticii* DNA fragments.

Synthesis of duplex oligonucleotide conversion adapters:

Each oligonucleotide used to form the duplex conversion adapters was synthesized by and obtained from Oligos ET Inc. (Wilsonville, OR). One strand (A strand) of each duplex conversion adapter contains the EcoR I cohesive end (AATT) at the 5′ terminus to the 10 mer core annealing sequence. Three lengths of the "A strand" (A1, A2, and A3) were synthesized by the addition of single cytosine residues between the 5′ end of the core sequence and 3′ end of the EcoR I cohesive end. Oligonucleotides complimentary to each length of the "A strand" core annealing sequences (14 mer = B1, 15mer = B2, 16mer = B3) were synthesized with Sau3A I, Mbo I or BamH I cohesive termini (GATC) added to the 5′ end of the "B strand". The duplex conversion adapters were formed by separately annealing "A strands" and "B strands" with matching lengths of complimentary core sequences. For this, a 0.5 A_{260} unit of each of the lyophilized oligonucleotide was dissolved in 120 µl of distilled water to obtain a 50 µM solution. Forty µl of each of these complimentary oligonucleotides (A1+B1, A2+B2, A3+B3) were mixed with 10 µl of 10× buffer (250 mM Tris, pH 8.0, 100 mM $MgCl_2$) and 10 µl of distilled water. These mixtures were heated separately to 95°C and slowly cooled (approximately one hour) to room temperature. This yielded a 20 µM solutions of 1, 2 and 3 types of adapters. At this point the three lengths of each duplex conversion adapters with identical cohesive ends were stored separately at -80°C for future use.

Ligation of adapters to *Ehrlichia risticii* DNA fragments:

Dried ethanol precipitate of Sau3A I *E. risticii* restriction fragments (6 µg) was resuspended in 45 µl of distilled water and was aliquoted in three equal parts. Next, 15 µl of preannealed adapters type 1, 2 and 3 were added to parts 1, 2 and 3 respectively to yield approximately a 10:1 molar ratio of adapter to the insert fragments. To each of these mixtures, 5.0 µl of 10× ligase buffer (500 mM Tris, pH 7.5, 70 mM $MgCl_2$, 10 mM DTT), 0.5 µl of 10 mM ATP, 13 µl of distilled water, and 1.5 µl (6 Weiss units) of T4 DNA ligase (Stratagene, La Jolla, CA) were added, mixed well and incubated at 15°C for six hours. After completion of ligation reaction the contents of the three Eppendorf tubes were mixed together in one tube and were placed in a 70°C water bath for 10 minutes to heat inactivated the ligase

enzyme. Subsequently the tubes were cooled on ice.

Phosphorylation of adapter modified insert DNA and removal of excess adapters:

Adapter modified insert DNA was prepared for ligation into λ-ZAP vector (Stratagene, La Jolla, CA) by phosphorylation of adapter 5' ends with T4 polynucleotide kinase (Promega Corporation, Madison, WI) and removal of excess adapters by spin column chromatography.

Following heat inactivation and cooling, 150 µl of reaction mixture were added to 20 µl of 10× T4 polynucleotide kinase buffer (500 mM Tris-HCL, pH 7.5, 100 mM $MgCl_2$, 50 mM DTT, 1.0 mM spermidine), 10 µl of 0.1 mM ATP, 1.0 µl of T4 polynucleotide kinase (10 units), and 19 µl of distilled water. The reaction mixture was incubated at 37°C for 30 minutes and reaction was terminated by single extraction with 1 volume of TE- saturated phenol, followed by three extractions of equal volume of chloroform:isomyl alcohol (24:1). The upper aqueous phase was transferred to a fresh tube and unligated adapters were efficiently removed with spin column chromatography.

The Sephacryl S-400 matrix, spin columns, wash tubes and collection tubes for column chromatography were obtained from Promega Corporation (Madison, WI). The chromatography columns were prepared according to the instruction of the Promega technical bulletin (# 067). Briefly, Sephacryl S-400 slurry was thoroughly mixed and 1.0 ml slurry was transferred to a spin column. The column tip was placed in the wash tubes and then the whole assembly was placed inside a large centrifuge tube (Falcon #25319) and centrifuged in a swing bucket rotor at 800×g for five minutes. The wash tube with fluid in it was discarded, and a second centrifugation was performed in the same manner to discarded any remaining fluid in the column. The phosphorylated reaction mixture with excess adapters was applied to the top of the gel bed of the prepared column and the column was placed into the collection tube. This whole assembly was then centrifuged in the same manner as described before in the column preparation step. The phosphorylated adapter modified insert DNA present in the eluant of the collection tube was then ethanol precipitated at -20°C overnight by adding 0.5 volume of 7.5 M ammonium acetate and 2.0 volumes of ethanol. The precipitated DNA was

pelleted by centrifugation at 4°C for 15 minutes and the invisible pellet was washed once with 70% alcohol prior to vacuum drying.

Ligation of insert DNA to λ-ZAP arms:

The adapter modified phosphorylated vacuum dried insert DNA pellet was resuspended in 6.0 µl of TE (10 mM Tris, pH 8.0, 0.1mM EDTA). The optimal vector:insert ratio for efficient ligation was obtained by aliquoting 2.5, 0.5 and 0.1 µl of the *E. risticii* insert DNA into three separate tubes. One µg of EcoR I digested and dephosphorylated λ-ZAP arms (Stratagene) was added to each of the tubes, followed by 1.0 µl of 10× ligase buffer, 0.1µl of 10 mM ATP, and distilled water to 9.0 µl. Then 1.0 µl of T4 DNA ligase (4 Weiss units, Stratagene) was added and the solution incubated at 15°C for six hours.

Packaging of recombinants λ-ZAP DNA:

In vitro packaging of λ-ZAP concatomers was done using the commercially available packaging mix (Gigapack II Gold, Stratagene). Two µl of concatamerized λ-ZAP recombinants were added to a freeze-thaw extract tube. To this, 15 µl of sonicated extract were added and mixed well. After a brief spin to pull the contents to the bottom, tubes were incubated at 22°C for two hours. This packaging reaction was stopped by adding 500 µl of SM buffer (50 mM Tris, pH 7.5, 100 mM NaCl, 8.0 mM $MgSO_4$, 0.01% gelatin) and 20 µl of chloroform. The reaction mixture was centrifuged at 1500×g for five minutes to pull down the debris. The supernatant was transferred to another tube and 25 µl chloroform was added to it. This recombinant λ-ZAP stock was stored at 4°C for future use in titrating and screening.

Titration and amplification of the recombinants:

Fifty ml of LB (Luria-Bertani) broth supplemented with 0.2% maltose and 10 mM $MgSO_4$ were inoculated with a single colony of XL1-Blue strain of *E. coli*. The cells were grown overnight at 37°C in a shaking incubator. Then next day the cells were centrifuged at 1000×g for 10 minutes, and resuspended in 25 ml of 10 mM $MgSO_4$ and stored on ice. A 10-

fold serial dilution of phage stock (packaging mix), up to 10^{-10}, was prepared in SM buffer and 10 µl aliquots from each dilution were mixed and incubated separately with 200 µl of above prepared host cells. Each mixture was incubated at 37°C for 15 minutes to absorb the phage on the surface of the host cells. Seventy ml of NZY top agar (0.75%) were equilibrated at 48°C in a water bath; then 350 µl of 250 mg/ml 5-bromo-4-chloro-3-indoyl-β-D-galactopyranoside (X-gal) and 105 µl of 0.5M isopropyl-β-D-thiogalactoside (IPTG) were added to it. Seven ml of this molten top agar were mixed separately with each dilution of phage-bacteria mixture and poured immediately onto a 100 mm petri dish. The plates were incubated at 37°C for 6 hours and stored at 4°C overnight for color development. The next day, the blue and clear plaques were counted to determine the titer and cloning efficiency of the packaged λ-ZAP.

To obtain a high titer library for storage, *in vitro* packaged recombinants were amplified by plating approximately 50,000 plate forming units (pfu) and incubating at 37°C for about 6 hours. When the plaques attained the size of about 0.5 mm, 10 ml of SM buffer were added to the plate and incubated overnight while shaking at 4°C. The suspension containing phage was extracted once with chloroform and stored in the presence of 0.3% chloroform.

Immunoscreening of recombinants for expression of variant *Ehrlichia risticii* antigens:

The variant recombinant clones were screened for expression of variant *E. risticii* antigens using rabbit and mouse antisera against the variant *E. risticii* strain. Before use, the rabbit antisera was exhaustively absorbed against *E. coli* and λ-ZAP protein components.

Production of antiserum against variant strain:

Hyperimmune sera to the variant strain of *E. risticii* were produced in rabbits. The first injection contained 80 µg and 320 µg of the purified organism, emulsified with Freund's adjuvant, administered by interdermal and intramascular routes, respectively. A second injection, administered two weeks later by the intramascular route, contained 200 µg of the purified *E. risticii* emulsified in Freund's incomplete adjuvant. At four and seven weeks

following the first injection rabbits were again injected intramuscularly with 200 μg of the organism only. One week after the final injection, sera were collected and pooled. Antisera from mice infected or immunized with *E. risticii* were obtained as according to known procedures.

Absorption of variant *Ehrlichia risticii* antiserum:

Variant *E. risticii* antisera were exhaustively absorbed with the lysates of *E. coli* strain XL1-Blue and λ-ZAP phage to remove any non- specific antibodies. An one liter culture of XL1-Blue transformed with pBluescript SK-phagemids (Stratagene) was grown in LB medium to an OD_{600} of 0.5 at 37°C, and IPTG was added to 10 mM final concentration. The cells were harvested by centrifugation at 11,000×g for 10 minutes and the cell pellet was resuspended in 20 ml of 10 mM Tris, pH 7.5, 1.0 mM phenylmethylsulfonyl fluoride (Sigma). About 15 ml of cell suspension were subjected to four 30 second cycles of sonication at 4°C. Next, Triton X-100 was added to 0.05% and the homogenate was incubated for 30 minutes on ice and then diluted in 30 ml Tris saline (10 mM Tris, pH 7.4, 150 mM NaCl), and stored at -70°C. This preparation of bacterial cell lysate was designated as sonic lysate. To the remaining 5.0 ml of original cell suspension, Laemmli sample buffer (62.5 mM Tris, pH 6.8, 2% SDS, 5% 2-mercaptoethanol, 10% glycerol) was added to 1×, heated to 100°C for five minutes, and diluted with 10 ml of Tris saline and stored at -70°C. This preparation of bacterial cell lysate was designated as SDS lysate.

A large scale preparation of λ-ZAP phage particles was produced according to Maniatis. One liter of XL1-Blue cells was grown up to the OD_{600} = 0.5, in LB media supplemented with 0.2% maltose and 10 mM $MgSO_4$. The culture was inoculated with 10^{10} pfu of phage particles and incubated at 37°C for an additional five to six hours, until the visible lysing of the bacterial cells was prominent as indicated by presence of cell debris. The lysed culture was further incubated for 10 minutes in presence of 20 ml of chloroform. Pancreatic DNAase-I and RNAase (Sigma), were added to this lysed culture to a final concentration of 1.0 μg/ml and a further incubation was performed for an additional 30 minutes at room temperature. To disperse the phage particles from the bacterial debris, 58.4

gm of solid NaCl were added and the lysate was incubated at 4°C overnight. The next day the bacterial debris was removed from this lysate by centrifugation at 11,000×g for 10 minutes and 100 gm of solid polyethylene glycol (PEG 8,000) were mixed into the supernatant. The mixture was incubated on ice water for 1 hour and the precipitated phage particles were recovered by centrifugation at 11,000×g for 10 minutes. The supernatant was discarded and the phage pellet was resuspended in 20 ml of Tris saline and added to the sonic lysate obtained earlier.

In separate polyethylene bags, ten 137 mm nitrocellulose circles (NCM, Schleicher & Schuell, Inc., Keene, NH) were incubated with sonic lysate and another five membranes were incubated with SDS lysate for two hours at room temperature on a shaker. The membranes were then washed five times with Tris saline for 15 minutes each and incubated overnight with casein solution (2% casein in 10 mM Tris, pH 7.5, 120 mM NaCl) at 4°C. Five ml of rabbit anti- *E. risticii* serum were diluted in 100 ml of casein solution and placed in a tray. Two NCM adsorbed with sonic lysate and one NCM adsorbed with SDS lysate were placed in the tray and incubated for two hours. The membranes were taken out, replaced with new sets of membranes and incubated as before. The process was repeated with all the membranes. The absorbed serum was aliquoted and stored at -70°C.

Immunoscreening the recombinants of variant strain:

Screening the λ-ZAP recombinants for expression of *E. risticii* antigens was done according to known procedures. *E. coli* strain XL1-Blue was used as a host cell to plate the library. A liquid culture was started from a single colony and grown overnight with vigorous shaking at 30°C in LB media supplemented with 0.2% maltose and 10 mM $MgSO_4$. The cells were centrifuged at 1000×g for 10 minutes then gently resuspended in 0.5 volumes of 10 mM $MgSO_4$. About 700 to 1000 pfu of the packaged λ-ZAP were mixed with 1.2 ml of above prepared XL1-Blue cells and incubated at 37°C for 18 minutes. Twenty one ml of molten NZY top agar (0.8%), prewarmed to 42°C were then added, mixed, and poured onto a 150 mm plate containing 1.5% NZY bottom agar and the agar was allowed to solidify at room temperature for 15 minutes. The plates were incubated at 37°C for four hours, until the

plaques were about one mm in size. Next, a 137 mm colony/plaque screen membrane (NEN® Research products, Boston, MA) was saturated with IPTG solution (10 mg/ml) and blotted dry on a filter paper. This membrane was carefully placed on the top agar and incubation was continued at 37°C for another three hours. The membrane was pierced asymmetrically at three places with an 18 gauze needle, peeled from the agar, and washed three times with Tris saline to remove the debris and bacteria. The plates were then stored at 4°C and the washed NEN membranes were blocked with casein solution at 4°C overnight. The next day, membranes were incubated in a 1:100 dilution of the absorbed *E. risticii* antisera for two hours at room temperature and washed twice in Tris saline with 0.05% Triton X-100, and once in Tris saline for 15 minutes each. The antisera treated membranes were incubated either with 2.0 µg/ml of alkaline phosphatase labeled goat anti-rabbit IgG or mouse anti-rabbit IgG (Kirkegaard and Perry) for one hour at room temperature. The membranes were consecutively washed three times in the same way described earlier in this procedure, followed by a final wash with 0.9% NaCl. Finally the membranes were treated with Fast Red and naphthol substrate solution for about 10 minutes and the reaction was stopped by washing the membrane in distilled water.

The pink immunoreactive spots corresponding to the recombinants expressing *E. risticii* antigens were aligned with the help of the needle marks and those positive plaques were picked up from the plates with the aid of a Pasteur pipette. The agar plugs containing the recombinant plaques were dispensed separately into 500 µl of SM buffer and the phage were allowed to diffuse out by vortexing and incubating vials at 4°C for two hours. Twenty µl of chloroform were also added separately in each vial before long term storage. Plaque purification of the recombinants was accomplished by two additional rounds of immunoscreening as above.

Identification of recombinant antigens of variant strain:

The identity of the recombinant antigen expressed in the clones of the λ-ZAP library was established by preparing monospecific antigen-selected antibodies and reacting this with the nitrocellulose strips containing electrophoretically separated *E. risticii* antigens.

Production of recombinant clone specific antibody:

The plaque purified λ-ZAP recombinants (10^5 pfu) were mixed separately with 1.2 ml of pre-prepared $MgSO_4$ treated XL1-Blue cells and incubated at 37°C for 18 minutes to absorb the phage on surface of bacteria. Each phage bacteria mixture was plated on a 150 mm petri dish as described above. After the plaques had attained the size of 1.0 mm, a 137 mm NCM saturated in IPTG solution (10 mg/ml) was overlaid on the top agar of the plate and incubated at 37°C for four hours. The NCM was reversed and incubation was continued for an additional three hours. After washing and blocking the unbound sites, as described above, the NCMs were incubated with the 1:100 dilution of antisera at 4°C overnight. The membranes were washed once with Tris saline, twice in Tris saline with 0.05% Triton X-100, and once in Tris saline for 15 minutes each. The membranes were then placed separately in polyethylene hybridization bags (BRL) and 10 ml of glycine buffer (0.2 M glycine, pH 2.8, 150 mM NaCl) were added to each bag. The bags were heat sealed and incubated at room temperature for one hour to elute the antibodies. The eluted antibodies were neutralized to pH 7.0 with 500 µl of 1.32 M Tris base and stored at -70°C. A preparation made from the non recombinant λ-ZAP was processed in same way as the negative control.

Identification of the recombinant antigens:

The recombinant clone-specific antibodies were diluted with an equal volume of casein solution. These antibodies were incubated overnight at 4°C with a strip of NCM on which electrophoretically separated *E. risticii* proteins had been blotted. Next, the strips were treated with alkaline phosphatase labeled anti-rabbit IgG and substrates. The strips were now aligned with an adjacent strip which had been reacted with polyclonal *E. risticii* antisera and the identity of the antigen encoded by the recombinant was ascertained.

Antigenic analysis of standard and variant strains:

Western immunoblotting was performed on both standard and variant *E. risticii* strains with their homologous and heterologus mouse antisera by the procedure described above. Antigenic analysis of the component antigens of these two different strains of *E.*

risticii was also performed by Western blotting with clone specific antibodies of 85kD, 55kD, 51kD and 28kD proteins of the variant strain and the 50kD of the standard strain. In order to perform this renograffin purified *E. risticii* of standard and variant strains were gel electrophoresed in several lanes in alternate combination. After blotting the NCM strips were cut in such a way that each strip contained the antigens of both the standard and variant strains. These strips were then treated separately with a clone specific antibody to determine the antigenic difference and similarities between the strains. The techniques which were followed here were described above.

Construction of DNA probe and DNA hybridization:

The random primer labeling technique was used to incorporate the radioactive ^{32}P in *E. risticii* insert DNA of several recombinants. The labeled probes generated in this manner were used in Southern hybridization of the restricted *E. risticii* genome of standard and variant strains.

Probe DNA:

The probe DNAs were prepared by using two different techniques. One of the techniques involved restriction digestion of the recombinant phagemids, agarose gel electrophoretic separation of the insert DNA, and elution of the insert DNA band from the gel. In this process, after gel electrophoresis, the agar piece containing the specific DNA band was visualized on an ultraviolet light transilluminator (Hoeffer, San Francisco, CA) and separated out from the gel by using a razor blade. The DNA was extracted from the agarose gel piece by using silica beads (GenecleanII, Bio101, La Jolla, CA), following the manufacturer's suggested protocol. Briefly the gel piece was weighed and about three volumes of 6.0 M sodium iodide were added and incubated at 55°C until the agarose completely dissolved. To this, 5.0 µl of silica beads (Glassmilk, Bio101) were added and the emulsion incubated at room temperature for five minutes while occasionally mixing the silica beads with the dissolved agarose by tapping the tube. The silica beads were separated out from the solution by quick centrifugation in a table top microcentrifuge and washed with 10

mM Tris, pH 7.5, 100 mM NaCl, 1.0mM EDTA, 50% ethanol (New Wash; Bio101). The process was repeated for two more times and finally the DNA bound with the silica beads was eluted by resuspending the beads with 10 µl of distilled water and incubating at 55°C for five minutes.

The other technique involved PCR amplification of a segment of DNA directly from the genomic DNA of *E. risticii*. Specific sequences obtained from the cloned *E. risticii* recombinants were used to select the proper primer pair for each amplification. Typically a 100 µl PCR reaction mixture consisted of 10 µl of 10× reaction buffer (0.5M KCl, 0.1M Tris, pH 8.3, 15 mM $MgCl_2$ and 0.1% gelatin), 16 µl of deoxyribonucleotide triphosphates (160 nmoles each), 4.0 µl of primers (0.1 nmole each), 0.5 µl of *Taq* polymerase (5units/µl), 10 µl of *E. risticii* genomic DNA (1 µg/ml) and 59.5 µl of distilled water. The amplification was performed using a DNA Thermal Cycler (Perkin Elmer Cetus, Norwalk, CT). The initial template denaturation step proceeded for 1.5 minutes at 95°C. Then a typical cycle profile consisted of annealing for two minutes at 52°C, extension for three minutes at 72°C and denaturation for one minute at 94°C. A total of 60 cycles were performed. At the end of the 60th cycle the heat denaturation step was omitted and the extension step was extended by an additional seven minutes. Following the termination of the amplification cycle, the sample was allowed to return to at 4°C temperature and held there. The specificity of the PCR amplified DNA was further confirmed by gel electrophoresis and the DNA was purified by Geneclean II, following the procedure described above.

Labeling of probe DNA:

The random primer labeling was done by Prime a Gene® Labeling System (Promega Corporation). Twenty ng of DNA were diluted in 25 µl of distilled water and boiled at 100°C for two minutes. The solution was immediately chilled on ice. With the tube held on ice, 10 µl of 5× labeling buffer (250 mM Tris-HCl, pH 8.0, 25 mM $MgCl_2$, 10 mM DTT, 1 M HEPES, pH 6.6, and 26 A_{260} units/ml random hexadeoxyribonucleotide primers), 2.0 µl mixture of nonlabeled deoxyribonucleotides (0.5 mM each of dATP, dGTP, and dTTP), 2.0 µl nuclease free BSA (400 µg/ml), 5.0 µl of $\alpha^{32}P$ (50 µCi) dCTP and 5.0 units of DNA

polymerase I Klenow fragment were added. The mixture was incubated overnight at room temperature and finally the labeling reaction was terminated by heating at 100°C for two minutes. The unincorporated nucleotides from the reaction mixture was removed by using the Push column (Stratagene) and the mixture was stored at -20°C for future use in a Southern hybridization cocktail.

The amount of incorporated radioactivity and the specific activity was measured by TCA precipitation (10% trichloroacetic acid and 1% sodium pyrophosphate). Specific activities in the range of 10^9 counts/minutes (CPM) per μg of DNA were obtained by the random primer labeling method.

Southern blot hybridization:

The gel containing the DNA samples was acid-depurinated with 0.25M HCl for 15 minutes, denatured with 0.4 M NaOH-0.6 M NaCl for 30 minutes, and neutralized with 1.5 M NaCl/0.5 M Tris, pH 7.5, for 30 minutes. In preparing for capillary transfer, two layers of 3.0 mm Whatman filter paper were spread on the Plexiglas support of the blotting apparatus (BRL), and placed in a buffer tray filled with 10× SSC and a pipette was rolled over them to remove air bubbles. The gel was carefully inverted and placed on the filter paper, and again the pipette was rolled over to remove any trapped air bubbles. This whole assembly was then covered with Saran-wrap and the plastic was cut exactly to the outline of the gel. The Saran-wrap used to avoid direct contact between the wick and the stacking paper towels, and it also helped to prevent the unnecessary evaporation of the buffer. Genescreen Plus Membranes (NEN products) were presoaked in distilled water for one minute and in 10× SSC for 15 minutes. The membranes were placed on the gels and air bubbles were removed by rolling the pipette as before. Three sheets of 3.0 mm Whatman filter paper were soaked in 10× SSC and placed on the blotting membrane and a 5" stack of paper towels were laid on the top of these filter papers. The DNAs from the gel were capillary transferred to the Genescreen Membrane overnight. The next day, the membrane was carefully removed from the gel and was treated with 0.4 N NaOH for one minute and 2.0× SSC/0.2 M Tris, pH 7.5 for five minutes. Finally the membrane was placed on a 3.0 mm Whatman filter paper and DNAs

were permanently bound with the membrane by a 30 seconds exposure in an Ultra Violet Crosslinker from Stratagene.

The membrane blot was prehybridized in a polypropylene bag for 1.5 hours at 45°C in a prehybridization solution consisting of 8 ml of Hybrisol I (Oncor, Gaithersburg, MD) and 2.0 ml of Hybrisol II (Oncor) solution. The probe DNA was denatured by boiling for two minutes and added to a final concentration of 3×10^6 counts/minutes (CPM) per ml of prehybridization solution. The hybridization was continued at 45°C for 18 hours and then membrane blot was carefully removed from the bag. The membrane was washed twice with 1.0× SSC, 0.1% SDS at room temperature for 20 minutes each and once with 0.1× SSC, 0.1% SDS at 60°C for one hour. The wet membrane was sealed in a hybridization bag and exposed to X-omat film (Kodak, Rochester, NY) at -70°C for varying time intervals. The multiple rehybridization of the same membrane blot was also accomplished by stripping the probe from the membrane. To do this the membrane was boiled for 30 minutes in a solution of 10 mM Tris-HCl, pH 8.0, 1 mM EDTA and 1% SDS. The DNA molecular weights in a Southern blot were determined by hybridizing the 1 kilobase DNA ladder (BRL) run in the adjacent lane of the gel with the ^{32}P labeled probe of one kilobase DNA ladder.

Recombinant DNA procedures and sequencing:

The variant *E risticii* antigens were expressed by several λ-ZAP recombinants. The *in vivo* excision of those λ-ZAP recombinants yielded pBluescript SK(-) phagemid clones. The specific clones obtained from the recombinants expressing 85kD antigen were further subcloned to obtain the complete nucleotide sequence of the 85kD gene. The λ-gt11 recombinant of 50kD antigen gene was cloned in pBluescript SK(+) phagemid.

In vivo excision of pBluescript SK(-) phagemid:

In vivo excision of the pBluescript SK(-) phagemids from the λ-ZAP recombinant phages was done according to the procedures of the manufacturer (Stratagene). The ExAssist™ helper phage and Solr™ bacterial strain [*e14-(mcrA), Δ(mcrCB-hsdSMR-mrr) 171, sbcC, recB, recJ, umuC::Tn5(kan^r), uvrC, lac, gyrA96, relA1, thi-1, endA1, λ^R, {F'*

proAB, lac,^{q} *Z M15)* Su^{-} (non-suppressing)] were also obtained from the Stratagene. After being plate purified three times, a single recombinant plaque was lifted from the agar plate and transferred into a sterile microfuge tube containing 500 µl of SM buffer and 20 µl of chloroform. The tube was vortexed and incubated at room temperature for two hours to diffuse the phage from the agar block into SM buffer. The titer of this phage stock was 10^6 pfu/ml. In a 50 ml tube, 200 µl of 0.5 OD_{600} XL1-Blue cells were mixed with 100 µl of phage stock and 1.0 µl of ExAssist helper phage and incubated at 37°C for 15 minutes. Next, 3.0 ml of 2 × YT media were added and incubation was continued for another 2.5 hours at 37°C in a shaker incubator. In this incubation period, a co-infection of the recombinant λ-ZAP phagemid and the ExAssist helper phage proceeded in the same XL1-Blue cells. As a final result the newly created recombinant pBluescript SK(-) phagemids were packed inside of the ExAssist helper phage and released from the bacterial cells. Once the phagemids were secreted, the remaining XL1-Blue cells were removed from the mixture by heating the tube at 70°C for 20 minutes. The heat treatment killed all the bacterial cells while the phagemid remained resistant to the heat treatment. The heat inactivated mixture was then centrifuged at 4,000×g for 10 minutes to pellet the cellular derbies and the supernatant was stored at 4°C.

To rescue the phagemid, 10 µl and 0.1 µl volumes of packaged phagemid stock from above were mixed with 200 µl of 0.1 OD_{600} Solr cells (*E. coli*) separately and incubated at 37°C for 15 minutes. About 10 to 50 µl volumes were plated onto LB plates containing 100 µg/ml of ampicillin, and incubated at 37°C overnight. Since the Solr cells were resistant to λ-ZAP recombinant, the colonies which appeared the next day on the plates contained the pBluscript SK(-) double stranded phagemid with the cloned DNA insert. The bacteria infected with helper phage alone could not grow because they did not contain the ampicillin resistant gene.

Extraction and purification of phagemid DNA:

Phagemid template DNA was prepared for sequencing and other recombinant work by a known method. The cell containing the phagemids were grown in lightly capped 15 ml plastic screw cap tubes with 5.0 ml of LB broth containing 100 µg/ml ampicillin. The

cultures were aerated by mixing them in a shaker incubator at 37°C overnight. The following day 1.5 ml of the cultures were transferred to 1.5 ml microfuge tubes and centrifuged for two minutes. The supernatant was removed, an additional 1.5 ml of culture was added, and the tubes were again centrifuged for two minutes. The supernatant was removed as completely as possible and cellular pellet was resuspended in 100 µl of an ice cold solution of glucose / Tris / EDTA buffer (50 mM glucose, 25 mM Tris-HCl, pH 8.0 and 10 mM EDTA). The tubes were incubated at room temperature for five minutes. Cells were lysed by the addition of 200 µl NaOH/SDS solution (0.2 N NaOH, 1% SDS), gentle mixing, and incubation at room temperature for 10 minutes. Neutralization of NaOH and precipitation of SDS and chromosomal DNA was accomplished by the addition of 150 µl of 3.0M potassium acetate, pH 4.8 with gentle mixing for at least 30 seconds. The contents were centrifuged for five minutes at 4°C and supernatants were transferred to fresh tubes, centrifuged another five minutes, and again transferred to new tubes, avoiding the carryover of any precipitate. To these supernatants, 1.0 ml of ice-cold absolute ethanol was added and the nucleic acids were allowed to precipitate at 20°C for 30 minutes. The nucleic acid precipitates were collected by centrifugation at 4°C for five minutes, washed with 70% ethanol, and the pellets were dried. The nucleic acids were resuspended in 20 µl of TE buffer and the RNA was digested by the addition of 1.0 µg of RNAase-A at 37°C for 30 minutes. About 2.0 µl of this mini preparation DNA were used for restriction digestion analysis.

Subcloning of 50kD recombinant of 25D strain:

The insert of 50kD antigen gene of 25D strain identified from the λ-gt11 library, was re-cloned in pBluescript SK(+) for restriction mapping. This insert- plasmid of recombinant pBluescript SK(+) was restriction digested and the fragments were subcloned in pBluescript SK(-) vector for further analysis and sequencing purposes. The internal segment of the insert was also PCR amplified and subcloned for the same interest.

The specific restriction digestion was obtained by the Hind III enzyme. For this, 1.0 µg of pBluescript SK(-) phagemid (Stratagene) was digested with Hind III, and the completeness of digestion was ascertained by agarose gel electrophpresis. The DNA was

then extracted with phenol:chloroform and resuspended in 1.0× calf intestinal alkaline phosphatase buffer (Promega, 50 mM Tris, pH 9.0, 10 mM $MgCl_2$, 1.0 mM $ZnCl_2$, 10 mM spermidine). Dephosphorylation of the 5' PO_4 groups was accomplished by digestion with 2.0 units of calf intestinal alkaline phosphatase (Promega). The enzyme was removed from the reaction mixture by phenol-chloroform extraction and DNA was ethanol precipitated as before, with additional washing in 70% ethanol to remove the pyrophosphate ions. Finally the DNA pellet was resuspended in 10 µl of TE buffer. About 2.0 µl of mini preparation DNA (1.0µg) were mixed with 4.0 µl of the appropriate 10 × digestion assay buffer (Promega) and Hind III restriction endonuclease (Promega) at a final concentration of 1.5 unit/µg DNA. After complete digestion for one hour at 37°C, 8.0 µl of the gel loading buffer (Appendix 5) containing a marker dye, were added to the tube. The reaction mixture was electrophoresed on 1% agarose gel by a submerged horizontal gel electrophoresis apparatus (BRL). Marker DNA (1 kilobase ladder, BRL) was electrophoresed simultaneously to monitor and compare the run of the DNA samples. Upon completion of the electrophoretic run, the migration pattern of the DNA bands was viewed with a 302 nanometer ultraviolet transilluminator (Spectoline, Model T. P.-302). The upper band consisted of plasmid DNA and the lower two bands consisted of insert DNA of the 50kD antigen gene. The insert bands, as ascertained by electrophoretic migration, were cut out from the gel and processed for purification of DNA by GenecleanII (Bio101) silica matrix.

One µl of the prepared vector (0.1µg) was mixed with two different 10 fold dilutions of insert DNA to obtain a nearly optimal ratio (2:1, insert:vector). To each of these reaction mixtures, 1.0 µl of 10 mM ATP and , 1.0 µl of 10 × ligase buffer (Promega, 1.0 × is 3.0 mM Tris, pH 7.8, 10 mM $MgCl_2$, 10 mM DTT and 5.0 mM ATP), were added. Each was brought to a final volume of 9.5 µl with distilled water. The DNA ends were ligated with two units of T4 DNA ligase (Promega) and the solution were incubated overnight at 18°C.

The *E. coli* XL1-blue competent cells were transformed with the ligated DNA by electro-transformation, using the Bio-Rad Gene Pulser apparatus. The competent cells were produced according the procedure described in Pulse controller instruction manual (Catalog = 165-2098) of Bio-Rad. One liter of LB broth was inoculated with 1/100 volume of a fresh

overnight culture and grown at 37°C with vigorous shaking to an OD_{600} of 0.6. The rapidly growing culture was cooled on ice for 30 minutes and the cells were harvested by centrifugation at 4,000×g for 15 minutes in 4°C. The pellet was washed two times with one liter of ice cold distilled water and finally the pellet was resuspended in 3.0 ml of 10% glycerol. The prepared cells were aliquoted and stored at -70°C. Just before the electro-transformation, the frozen cells were thawed on ice and 40 µl of the cell suspension were added to 2.0 µl of ligation mix. After one minute incubation on ice, the mixture was transferred into a cold 0.2 cm electroporation cuvette and pulsed with a time constant of four to five milli seconds with a field strength of 12.5 kV/cm. Immediately 1.0 ml of prewarmed SOC medium (2% Bacto tryptone, 0.5% Bacto yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose) was added to the mixture and then incubated at 37°C for one hour in an orbital shaker. About 20 µl aliquots were plated on LB plates containing 100 µg/ml ampicillin, X-gal, and IPTG. After overnight incubation at 37°C, single white recombinant colonies were picked up, grown in 5.0 ml of LB medium with 100 µg/ml of ampicillin and stored at -70°C in 15% glycerol.

Due to the presence of the direct repeats, the central segment of the 50 kD antigen gene was PCR amplified and subcloned separately. This strategy was followed to avoid the binding of sequencing primer at more than one place in the entire length of the gene. To do this, 1.0 pg of 50 kD recombinant plasmid DNA was PCR amplified following the standard reaction parameter described above. The amplified product was ascertained by electrophoretic migration and the product was cloned in pCR™ II vector (Invitrogen). Briefly 1.0 µl of PCR amplified product was mixed with 5.0 µl of distilled water, 1.0 µl of 10× ligation buffer, 2.0 µl pCR™ II vector and 4 units of T4 DNA ligase. The mixture was incubated at 12°C overnight. *E. coli* One Shot™ competent cells (Invitrogen) were transformed with the ligated DNA according to the manufacturer's procedure. Fifty µl of competent cells were thawed on ice and 2.0 µl of 0.5 β-mercaptoethanol and 1.0 µl of ligation mix were added to it. After incubation on ice for 30 minutes, the cells were heat shocked by placing in a 42°C water bath for 30 seconds and immediately transferred to ice for two minutes. The transformed cells were re-vitalized by adding 450 µl of pre warmed SOC media

and shaking in a incubator at 37°C for one hour. About 100 µl aliquots were plated on LB plates containing 50 µg/ml ampicillin and X-gal. After incubation at 37°C overnight, single white colonies were picked up, grown in 5.0 ml of LB medium with 100 µg/ml of ampicillin and stored at -70°C in 15% glycerol.

Subcloning of 85kD recombinants of 90-12 strain:

Two different recombinant phagemid clones expressing the 85kD antigen gene were identified from λ-ZAP library of 90-12 strain. After *in vivo* excision the recombinant phagemids DNA were extracted and the size of the inserts from these two specific clones were ascertained by Hind III and Sau3A I restriction enzyme digestion. Several insert fragments from Sau3A I restriction digestion products were further subcloned in pBluescript SK(-) vector for sequencing purposes. The two clones which expressed the 85kD antigen gene did not cover the 5′ end of this gene. To clone the 5′ end and obtain the complete sequence, the upstream 5′ region of the 85kD gene was PCR amplified directly from the genomic DNA of 90-12 strain. The specific primers used for this purpose were selected from the 5′ upstream and middle of the 50kD gene sequence of 25D strain and the PCR product was cloned in pCR™ II vector. This was accomplished according to the procedure described above.

DNA sequencing:

The double stranded DNA was sequenced according to the Sangers dideoxy chain termination method using the Sequenase® Version 2.0 kit (United States Biochemical, Cleveland, Ohio). This method involved the *in vitro* synthesis of a DNA strand from a single stranded DNA template using a DNA polymerase. Synthesis was initiated at only one site where an oligonucleotide primer annealed to the template. The synthesis chain reaction was terminated by the incorporation of a nucleotide analogue that would not support continued DNA elongation (hence the name chain termination). The chain terminating nucleotide analogues were the 2′, 3′ dideoxynucleoside 5′-triphosphates (dd NTPs) which lacked the 3′-OH group necessary for DNA chain elongation. When proper mixtures of dNTPs and one of

the four ddNTPs were used, enzyme catalyzed polymerization was terminated in a fraction of the chain population at each site where the ddNTPs were incorporated. Four separate reactions, each with different ddNTPs, gave complete sequence information. A radiolabeled nucleotide was incorporated during the synthesis, so that the labeled chain of various lengths were visualized by autoradiography, after separation by high resolution electrophoresis.

The polymerase 'Sequenase®' a modification of bacteriophage T7 DNA polymerase (United States Biochemical), was used for sequencing. The unique properties of Sequenase® are high processivity, low 3′ to 5′ exonuclease activity, and the efficient use of nucleotide analogues. These characteristics produce radioactive bands of more uniform intensity and less background radioactivity than those obtained when using a large fragment of *E. coli* DNA polymerase I or reverse transcriptase. Synthetic oligonucleotides (Oligos ETC Inc), specific for DNA clones at different restriction sites, were used as sequencing primers. Template DNA, purified by minipreparation was first annealed to the sequencing primer. Then DNA synthesis was carried out in two steps. The first step labeling and the second step resulted in the accurate termination of DNA synthesis using the dideoxynucleotides. In the first step, the primer was extended using a limiting concentration of deoxynucleoside triphosphates, including the radiolabeled dATP. In this step, virtually complete incorporation of labeled nucleotide occurred into DNA chains which were distributed randomly in length, from several to hundreds of nucleotides. In the second step, the concentration of all the deoxynucleoside triphosphates were increased and a dideoxynucleoside triphosphate was added. Processive DNA synthesis occurred until all growing chains were terminated by a dideoxynucleotide. At this stage, the chains were extended on an average of several dozen nucleotides. The reaction was ultimately terminated by the addition of EDTA and formamide. This was followed by denaturation electrophoresis and autoradiography.

Annealing of template and primer:

The miniprep, RNA-free double stranded plasmid DNA was first denatured by the alkaline denaturation method prior to annealing the sequencing primer with the target sequence. To do so, 8.0 µl of miniprep DNA was mixed with 9.0 µl of distilled water, 2.0 µl

of 2M NaOH, 1.0 µl of 4.0 mM EDTA, and the mixture was incubated at 37°C for 30 minutes in a water bath. The mixture was neutralized by adding 0.1 volume of 3 M sodium acetate (pH 5.0) and the DNA was precipitated with three volumes of ethanol at -70°C for 15 minutes. After washing the pelleted DNA with 70% ethanol, it was redissolved in 7.0 µl of distilled water and 2.0 µl of Sequenase® (United State Biochemicals) reaction buffer, and 1.0 µl (3.0 ng) of the appropriate primer was added. The mixture was heated to 65°C for two minutes and then slowly cooled down to ambient temperature over a period of 30 minutes. Once the temperature was below 35°C, annealing was complete.

Labeling reaction:

To label the DNA, a labeling mix, (supplied with the kit) was diluted five fold with distilled water (2.0 µl of labeling mix and 8.0 µl of distilled water) in a sterile Eppendorf tube. One µl of Sequenase was diluted with 7.0 µl of ice cold TE buffer in another sterile Eppendorf tube. To the Eppendorf tube containing 10 µl of annealed template-primer, the following were added sequentially: 1.0 µl of 0.1M dithiothreitol, 3.0 µl of diluted labeling mix, 0.5 µl of dATP (10/µci/µl), and 2.0 µl of diluted Sequenase®. After mixing, the tube contents were incubated for five minutes at room temperature.

Termination reaction:

Four Eppendorf tubes were labeled A, C, G and T. Two µl of each termination mix (supplied in Sequenase® kit) were placed in the respective tubes. The termination tubes were prewarmed to 37°C for one minute in a water bath. When the labeling reaction was completed, 3.5 µl of labeling mixture was transferred into each termination tube. The contents were mixed and incubated at 37°C for 5 minutes in a water bath. Following incubation, 4.0 µl of stop solution (95% formamide, 20 mM EDTA, 0.05% bromophenol blue, 0.05% xylene cyanol) were added to each tube to stop the reaction. The contents of the tubes were mixed thoroughly and stored at -20°C until ready to load on the sequencing gel for electrophoresis.

Sequencing gel electrophoresis:

A Baserunner 200 Sequencing apparatus (Eastman Kodak Company, Rochester, NY) was used for electrophoresis of the sequencing gel. To cast the 6% polyacrylamide gel, two clean, Sigmacote (Sigma) treated glass plates were assembled using a vinyl side spacer (0.4 mm) and 50 ml gel mixture (28.35 g urea, 10.5 ml 40% bisacrylamide, 6.75 ml 10× TBE, 26 ml of distilled water, 675 µl of 10% ammonium persulfate, 18 µl of TEMED) was poured into the gel mold. The flat edge of the shark-tooth comb (0.4 mm) was inserted between the plates to a minimum depth of 2.0 to 3.0 mm. After overnight polymerization, the comb was removed and then placed again with it's teeth facing the gel sandwich.

Then each buffer chamber of the apparatus was filled with approximately 500 ml of electrophoresis buffer (1× TBE). The gel was pre-electrophoresed for 30 minutes at a constant power of 60 watts, before loading the samples. The DNA samples from the dideoxy sequencing reactions were heated to 80°C for two minutes and then transferred to ice immediately prior to loading onto gel. The wells of the gel were rinsed out using a 10 ml syringe, attached with a 18 gauge needle to remove urea that had diffused out from the gel. Three µl of sample from each tube marked A, C, G and T were loaded onto the gel in the wells in that order (left to right). After loading, the sequencing gel was electrophoresed at 55 watts to generate enough heat to keep the DNA denatured. The surface temperature of the glass plate was maintained at least 50°C during electrophoresis. About three hours later, when the lower marker dye reached the bottom of the gel, another 3.0 µl of each sample were loaded into new wells in the same order and the gel was electrophoresed at 52 watts for another two hours. After the samples were run, the upper glass plates were disassembled carefully and the gel was soaked with 10% acetic acid and 12% ethanol until the xylene cyanol disappeared. This was done to ensure that all the urea was removed from the gel. The gel was removed from the lower glass plate onto a support of 3.0 mm Whatman paper and placed in a gel dryer for two hours.

The dried gel was placed in a metal cassette which had a spring-loaded lid to hold the gel and the film in a close contact. The gel was exposed to X-Omat™ (Eastman Kodak Company) 18×43 cm film in direct contact with the gel. After overnight exposure, the film

was removed and developed by using an automatic X-ray developer.

Analysis of DNA and deduced amino acid sequences:

The DNA sequence analysis was done by IBI Pustell software (IBI Limited, Cambridge, England). Using the program "Protein Coding Region Locator" the open reding frame (ORF) of DNA sequences were acertained. This program combines several method for locating potential coding regions in a DNA sequence. The first method searches both strands of the DNA sequence, looking for regions between user-set start and stop codons (ORF). In prokaryotes it uses ATG for starts and termination codons for stops, and searches for all possible six reading frames. The second method uses a statistical search (Fickett's Testcode) which looks for regions of DNA with biased usage of codons. This measurement is made over a window of bases which Fickett has shown must be at least 200 for good results. The probability can be set (0.29, 0.40, 0.77 or 0.92) to confirm that the region located is a real coding region. The high value of 0.92 was used for this analysis to maintain a high stringency condition. A combined test was performed to get a potential region meeting both criteria.

The amino acid sequence analysis was performed by using Peptide-Structure and Plot-Structure programs (PepPlot). PepPlot was written by Drs. Michael Gribskov and John Devereux of the Genetics Computer Group, and it was available through National Institute of Health (NIH, Bethesda, MD). Peptide-Structure makes secondary structure prediction for a peptide sequence. The predictions measure for antigenicity, flexibility, hydrophobicity and surface probability. Plot-Structure displays these predictions graphically. Using this program the secondary structure of a protein was predicted according to the Chou-Fasman method hydrophilicity according to the Kyte-Doolittle method and antigenic index according to the Jameson-Wolf method.

Expression of 50kD and 85kD homologue antigen genes:

After the full sequence analyses of the 50kD and 85kD major antigen genes, they were cloned separately by PCR in the expression vector pRSET C (Invitrogen). The advantage of

using this expression system was that the foreign prokaryotic genes were expressed in high amounts by the bacteriophage T7 promoter present upstream of the cloned genes. This high level of expression was facilitated by infecting the *E. coli* cells with M13 phage which expressed T7 RNA polymerase. For the cloning of the 50kD and 85kD antigen genes, two primers, one at the 5′ end of start site of the gene and the other at the 3′ end of the termination site of the gene, were selected. While synthesizing these primers, a sequence containing one restriction enzyme site was added eight bases upstream of the 3′ end of each primer. Different restriction enzyme sites were added to the two primers so that the amplified product could be cloned in the desired orientation in the multiple cloning site of the vector. The expression of the cloned gene and purification of the expressed protein was done according to the recommendations of the manufacturer (Invitrogen). For easy purification of the expressed protein, a metal binding polyhistidine domain and a site for enterokinase cleavage were also added by the vector sequence to the amino-terminal of the recombinant protein. The expressed recombinant protein could be purified by binding to a nickel (Ni^{2+}) charged resin and the extra Ni^{2+} domain could be cleaved using enterokinase.

PCR amplification of 50kD and 85kD genes:

The complete 50kD and 85kD genes were amplified separately from the genomic DNA of the original and variant strain of *E. risticii* by using two modified primers, named as expression cloning primers E.C.P-1 and E.C.P-2 (Figure 1). The E.C.P-1 (5′ CAT AAA ATT TCT AAG ACG AAG GAT CCC TAT GTC 3′) was selected from the known sequence of bp upstream of the first methionine codon of the genes. This 33 base primer was modified at base 21 and 22 position by the substitution of two A's in it original sequence with two G's. In the same way, E.C.P-2 (5′ GAC AGA AAG TTC CCC GTG TGA ATT CTA GCT AGG 3′) was selected from the known sequence 69bp downstream of the stop codon of the gene. This 33 base primer also was modified at base 21 by introducing another single base A. Amplification of the complete genes (50kD and 85kD) by using these two modified primers produced BamH I and EcoR I sites at the extreme 5′ and 3′ end of the genes respectively.

PCR amplification was accomplished according to the standard protocol described above. As a template, the genomic DNA of the original and variant strains were produced directly from their respective cell culture materials by the PCR lysis method (9). *E. risticii* infected HH cells (1 million) were harvested on day five to seven postinfection and the cell pellet was frozen and thawed three times to rupture the cells. Then 1.0ml of PCR lysing buffer [50 mM KCl, 10 mM Tris (pH 8.3), 2.5 mM $MgCl_2$, 0.1 mg/ml gelatin, 0.45% Nonidet P40 (Sigma Chemicals), 0.45% Tween 20 and 0.06 mg/ml K] was added to the ruptured cells and the mixture was incubated at 62°C for one hour. Finally the mixture was incubated at 95°C for seven minutes to heat inactivate the protinase K and stored at 4°C. Ten µl of this preparation were used as a template for PCR amplification.

Cloning the PCR amplified products in SK(-) phagemid:

The PCR amplified 50kD and 85kD genes were cloned separately in the multiple cloning region of the Sk(-) phagemid (Stratagene) for further sequence analyse. This extra step of cloning and sequencing was accomplished to confirm the correct amplification of the ORF prior to directional cloning in the expression vector pRSET-C.

To accomplish this amplified product (100 µl) was electrophorsed in an agarose gel and the specific DNA band was purified by Geneclean II, following the procedure described above. Fifty µl of this purified DNA were mixed with 6.0 µl of NEBuffer three (New England Biolabs), 3.0 µl of distilled water, and 10 units each of BamH I and EcoR I enzymes (New England Biolabs), and double digested at 37°C for one hour. The digested product was further purified by Geneclean II and eluted in 25 µl of distilled water. The SK(-) phagemid DNA (1.0 µg) was also double digested and purified separately by using the same technique and eluted in 25 µl of distilled water. The ligation of vector and insert DNA was accomplished by mixing 2.0 µl of double digested phagemid DNA with 4.0 µl of double digested PCR amplified product, 1.0 µl of 10× ligase buffer, 0.1 µl of 10 mM ATP, 1 µl (4 Weiss units) of ligase (Stratagene), and 1.9 µl of distilled water and incubating the mixture at room temperature for 2.5 hours. After the ligation reaction, the *E. coli* XL1-blue competent cells were transformed with 2.0 µl of the ligation mixture by electro-

transformation technique and the recombinant clones were selected by plating the transformed cells on LB plates containing 100 µg/ml ampicillin, X-gal, and IPTG. DNA from the recombinant clones were extracted by the minipreparation technique and sequenced to confirm the complete ORF of the 50kD and 85kD genes.

Cloning the 50kD and 85kD gene in pRSET-C expression vector:

Insert DNAs from the SK(-) recombinant clones of the 50kD and 85kD genes, which were confirmed by the sequence analyse, were used for further cloning in pRSET-C expression vector. The selected recombinants were grown in 50 ml of 2× YT media containing 100 µg/ml of ampicillin and recombinant phagemid DNA was extracted by minipreparation following the technique described above. For cloning, 1.0 µg of each pRSET-C plasmid DNA (Invitrogen) and recombinant phagemid DNA were taken separately and double digested with BamH I and EcoR I restriction enzymes in a 60 µl reaction volume. The reaction conditions were the same as described above. After the restriction digestion, both recombinant phagemid and pRSET-C plasmid DNA were electrophoresed in an agarose gel in separate lanes. The specific insert DNA bands and linear pRSET-C plasmid DNA bands were purified from the agarose gel by Geneclean II and eluted separately in 10 µl distilled water. One µl of the prepared vector (0.1 µg of pRSET-C) was mixed with two different 10 fold dilutions of insert DNA to obtain a nearly optimal ratio (2:1::insert:vector). To each of these reaction mixtures, 1.0µl of 10 mM ATP, 1.0 µl of 10× ligase buffer (Promega), and distilled water were added to a final volume of 9.5 µl. The DNA ends were ligated by incubation at room temperature for 2.5 hours in presence of 2 units (0.5 µl) of T4 DNA ligase (Promega).

The *E. coli* JM 109 [*recA1, supE44, endA1, hsdR17, gyrA96, relA1, thi*Δ(*lac-proAB*) F'(*traD36, proAB*$^{-}$, *lacI*^{q}, *lacZ*Δ*M15*)] competent cells were transformed with the ligated DNA by electro-transformation, using the Bio-Rad Gene Pulser apparatus. The competent cells were produced according the procedure described above and 40 µl of cell suspension were added to 2.0 µl of ligation mix. After incubation on ice for one minute, the mixture was transferred into a cold 0.2 cm electroporation cuvette and pulsed with a time constant of four

to five milliseconds with a field strength of 12.5 kV/cm. Immediately thereafter 1.0 ml of prewarmed SOC medium was added to the mixture and incubated at 37°C for one hour in an orbital shaker. About 50 µl aliquots were plated on SOB (2% Bacto trypton, 0.5% Bacto yeast extract, 8.5 mM NaCl, 2.5 mM KCl) plates containing 100 µg/ml ampicillin, X-gal and IPTG. After an overnight incubation at 37°C, single white recombinant colonies were picked up, grown in 5.0 ml of SOB medium with 50 µg/ml of ampicillin and stored at -70°C in 15% glycerol. Prior to the long term storage, further confirmation of the recombinants was ascertained by BamH I and EcoR I restriction digestion and agarose gel electrophoresis of miniprep DNA from +ve clones.

Expression of recombinant proteins in pRSET-C:

Each recombinant protein has different characteristics which can affect optimal expression parameters. To overcome this situation a pilot expression experiment was performed to determine the kinetics of induction for the 50kD and 85kD antigen genes. Briefly 2.0 ml of SOB media with 50 µg/ml of ampicillin were inoculated with a single white recombinant *E. coli* colony. The cells were grown at 37°C overnight in an orbital shaker. The next day 50 ml of SOB media with 50 µg/ml ampicillin was inoculated with 0.2 ml of the overnight culture and grown at 37°C with vigorous shaking to an $OD_{600} = 0.3$. An one ml aliquot of the culture was removed at this time point and centrifuged to pellet the cells. This was considered as the time zero sample and was frozen at -20°C. IPTG was added to the remaining culture to a final concentration of 1.0 mM and the cells were grown in presence of IPTG for an additional hour. After this time period the culture was inoculated with M13/T7 phage (Invitrogen) at an optimal ratio of 5 pfu/cell. The infection was allowed to proceed for another five hours at 37°C and an one ml aliquot of culture was removed every hour. Each sample was centrifuged and both the supernatant and cell pellet was stored as before.

After all the samples were collected, the each pellet was resuspended in 100 µl of 20 mM phosphate buffer (pH 7.0) and frozen in liquid nitrogen. The frozen samples were thawed again in a 42°C water bath and this freeze/thaw cycle was repeated an additional three times. Finally the freeze/thaw pellets were centrifuged at 14000×g for 10 minutes in a

refrigerated microcentrifuge and the supernatants with the soluble protein fractions were transferred to a fresh tube. The pellets with the insoluble protein fractions were also collected and resuspended in 100 µl of Laemmli buffer. The supernatants were also mixed with equal volumes of Laemmli buffer. Twenty µl of each sample (fractions of both supernatants and pellets) was electrophoresed separately on a 10% SDS polyacrylamide gel, following the identical procedures as described above. The gels were stained with Coomassie Blue and the bands were compared for increasing intensity in the expected size range of the 50kD and 85kD antigens to determine the optimal time point of maximum expression.

The large scale extraction and purification of the recombinant proteins were accomplished under denaturing condition. To do this, 50 ml culture of the selected bacteria expressing the recombinant proteins were harvested at the optimal time point of maximum expression. The cells were pelleted by centrifugation at 5,000 rpm for five minutes in a Sorvall SS-34 rotor and the pellets were resuspended in 10 ml of guanidine lysis buffer (6 M guanidine-HCl, 20 mM $NaPO_4$, 500 mM NaCl). The temperature of the buffer was preadjusted to 37°C for quick lysis of the cells, but to assure that complete lysis was obtained, the cells were rocked at room temperature for additional 10 minutes. To shear the DNA and RNA, the cell lysates were sonicated on ice with three five seconds pulses at a high intensity setting. After the sonication, the insoluble debris were removed from the sheared lysates by centrifugation at 3,000×g for 15 minutes and the clear lysates were stored at -20°C for further purification with ProBond™ resin columns (Invitrogen).

The recombinant proteins expressed in the pRSET-C vector contained six tandem histidine residues in the amino terminal of the peptides, which had a high affinity for ProBond™ resin. To bind the recombinant proteins in the columns, the resins of the columns were resuspended with 5.0 ml of guanidine lysate of the expressed proteins and rocked on an orbital shaker for 10 minutes at room temperature. The resins were settled by gravity and supernatants were removed carefully. This step was repeated again with another 5.0 ml fresh aliquot of the lysates. After binding the proteins with the resins, the columns were washed twice with denaturing binding buffer (8M urea, 20 mM $NaPO_4$, 500 mM NaCl, pH 7.8), twice with denaturing wash buffer (8M urea, 20 mM $NaPO_4$, 500 mM NaCl, pH 6.0) and twice with

the same denaturing wash buffer at pH 5.3. The washings were accomplished by simply resuspending the resins with 4.0 ml of each buffer for two minutes and then separating the resins from the supernatants by gravity. Finally the washed columns were clamped in a vertical position and the cap was snapped off on the lower end. The proteins were eluted from the columns by applying 5.0 ml of denaturing elution buffer (8M urea, 20 mM $NaPO_4$, 500 mM NaCl, pH 4.0). The elutes were collected and dialyzed against 10 mM Tris, pH 8.0, 0.1% Triton X-100 overnight at 4°C to remove urea, and then analyzed by Western blotting to confirm the specificity of the expressed proteins.

Immunoblot analysis of *Ehrlichia risticii* component antigens of 25D and 90-12 strains:

The antigenic composition profile of the variant (90-12) strain by Western blotting differed considerably from that of the 25D strain. Previous analysis of the Renografin purified standard (25D strain) indicated the presence of 18 component antigens of which nine (with molecular weights of 110, 70, 68, 55, 51, 50, 33, 28, and 22 kD) were major antigens. I^{125} surface labeling determined that the above antigens were apparent surface antigens. Further analysis by Western blotting with horse, rabbit and mouse antisera confirmed them as major antigens. Though several of these major antigens, namely the 68, 55, 49, and 28kD, proteins were similar in both strains, the main differences between them were as follows: (i) The 110 and 70kD antigens were present only in the 25D strain and they did not react with the 90-12 strain antisera. (ii) The 85kD antigen was present only in the 90-12 strain, but it reacted with the 25D strain antisera. (iii) The 50kD antigen was present only in the 25D strain and cross reacted with 90-12 strain antisera. (iv) The 55 and 51kD antigen bands in the 25D strain were well separated, whereas in the 90-12 strain they were close together as a 55/51kD band. (v) The 33kD antigen band of each strain showed comparatively less color intensity, with the heterologous antiseras as compared to the homologous antisera.

The recombinant antigens and their identity:

The recombinant clones expressing the partial or complete antigen genes were identified from two different genomic library of *E. risticii* strains. A λ-gt11 recombinant

library was constructed with 25D strain genomic DNA and a λ-ZAP recombinant library was constructed with 90-12 strain genomic DNA.

λ-gt11 recombinants:

The recombinant clones expressing the 50kD and 70kD antigen genes of 25D strain was produced previously in λ-gt11 bacteriophage. The 70kD recombinant was obtained from a library generated by using the partial Hpa II digest of *E. risticii* DNA. The 50kD recombinant was obtained from the library generated from *E. risticii* DNA subjected to a complete double digestion with Hpa II and HinP I. After identification of the recombinant antigens by the corresponding clone-specific antibodies, further analysis was conducted on 50kD antigen gene as a part of this study.

λ-ZAP recombinants:

The genomic expression library in λ-ZAP was generated with Sau3A I digested *E. risticii* DNA. The Sau3A I digested fragments ranged in size from about 400 bp to 2 kb. The efficiency of production of λ-ZAP recombinants was about 10^7 pfu/μg of λ-ZAP DNA, of which about 8% were non-recombinants. The number of antibody reactive recombinants was about 10 to 14 per 10^4 pfu. A total of 170 clones reactive with the 90-12 strain antisera were picked up for further analysis. Clone-specific antigen selected antibodies from these clones were prepared and reacted with strips of transblotted 90-12 strain antigens. A comparison of the Western blots of these clone specific antigen selected antibodies with polyclonal 90-12 antisera resulted in the identification of recombinants expressing the 85, 68, 55, 49, 33, 28, and 22kD antigens. The 51kD recombinant clone was identified separately from an EcoR I library of 90-12 strain generated in λ-ZAP system. After identification of recombinant antigens from this λ-ZAP library, further study was conducted on the 85kD antigen.

The expression characteristics and cross reactivity of 50kD and 85kD recombinant antigens:

A single recombinant clone expressing the 50kD recombinant antigen was identified from the λ-gt 11 library of the 25D strain, whereas two different recombinant clones expressing the 85kD recombinant antigen were isolated from λ-ZAP library of the 90-12 strain. Among these two recombinant antigens, the 50kD antigen was a nonfusion protein, expressed independently of the IPTG induction. Further analysis also revealed that the molecular mass of 50kD recombinant was identical to that of its native counterpart, indicating expression of the complete protein. Both 85kD recombinant clones expressed a partial 85kD antigen with β-galactosidase fusion. They migrated in the gel in conjunction with β-galactosidase, and their expression was dependent on IPTG induction.

It was discussed above that 85kD and 50kD antigens were not present in the 25D and 90-12 strains respectively, but these two proteins were cross reacted with each other's strain specific antisera. Further analysis by Western blot also revealed that the recombinant clone expressing only the 50kD antigen were cross- reactive with antisera raised in mice specifically against the 90-12 strain and the same way *vice-versa* with 85kD recombinant clones. These observations clearly indicated that the 50kD and 85kD antigens had their corresponding cross-reactive counter part present in both strains. As an attempt to identify these corresponding crossreactive counter parts of the two proteins in each strain, a Western immunoblot was performed with the 50kD and 85kD clone specific antibodies. It was observed that the 50kD clone specific antibody cross reacted with the 85kD antigen of the 90-12 strain and the 85kD clone specific antibody cross reacted with the 50kD antigen of the 25D strain. These results indicated the presence of common cross-reactive epitopes in two different molecular weight proteins which were very strain specific and distinguishable serologically. These two homologous proteins were designated as strain-specific antigens (SSA).

Nucleotide sequence of 50kD and 85kD recombinant clones:

The complete sequence of the 50kD antigen gene reading frame was obtained from a single clone identified in the λ-gt11 library of the 25D strain where as a partial reading frame of the 85kD antigen gene was obtained from two separate overlapping clones identified in λ-

ZAP library of the 90-12 strain. The remaining sequence at the 5' terminus of this gene was obtained later from a PCR amplified segment of the 90-12 genomic DNA. Both the 50kD and 85kD insert pieces were subcloned several times to obtain nucleotide sequence information and identify the possible open reading frame of both genes.

Subclones of 50kD and 85kD recombinants:

EcoR I restriction digestion of the 50kD λ-gt11 recombinant phage DNA, generated a 3.9 kb insert DNA fragment which was cloned in pBluscript SK(+) phagemid for restriction mapping. Fifteen restriction enzymes (6 base-cutters) were used to determine the presence of restriction sites in the insert DNA of the above pBluscript SK(+) subclone. The Hind III digestion of the recombinant pBluscript SK(+) DNA produced three DNA fragments of 3.5 kb, 2.2 kb and 1135 bp. The 3.5 kb DNA fragment was a plasmid-insert DNA piece, where 565 bp was an insert part and the rest of it was pBluscript SK(+) phagemid. This specific fragment was re-circularized to form a pB50-6.1 subclone. The 2.2 kb and 1135 bp insert fragments were subcloned separately in pBluscript SK(-) phagemid and they were designated as pB50-6.2 and pB50-6.3 respectively. It was difficult to select a primer for downstream sequencing of the pB50-6.2 recombinant clone, due to the presence of direct repeats in the middle of the insert. In order to overcome this situation an internal segment of 826 bp was PCR amplified by using two unique primers: 50-A ($^{5'}$ATA CTA AAA AGC ATA CTC$^{3'}$) and 50-B ($^{5'}$ TTC TAC AAG CCC TTT AAA$^{3'}$). The amplified product was cloned in pCR™ vector and designated as pCR50-6.2.1. The insert piece of the pCR50-6.2.1 recombinant clone was then easily sequenced by using the universal primers of the vector. The presence of direct repeat motifs in the pB50-6.3 recombinant clone produced the same problem as described above and thus the insert piece of this clone was further subcloned in smaller fragments to exploit the advantage of the universal primer sequences for the vector. For this purpose the restriction digestion was performed with Pst I and the generated fragments were cloned separately in pBluscript SK(-) phagemids. Subclones were designated as pB50-6.3.1 and pB50-6.3.2.

The two *in vivo* excised phagemid clones partially expressing the 85kD antigen gene

were designated as pB85-11 and pB85-17. The insert size of these two clones were 4.5 kB and 1.1 kb respectively. These two clones had 58% overlapping regions with each other and they together covered 84% of the 85kD gene sequence. The remaining unknown 16% of the 5' region of the gene was separately cloned by PCR from 90-12 genomic DNA, using primers 50-C ($^{5'}$GAA TGT TCA GCT TTC CGG$^{3'}$) and 50-D ($^{5'}$AGC TGT ATC GTT CGT GAG$^{3'}$). The 1.5 kb amplified product was cloned in pCR™ II vector and designated as pCR85-3. The 3' region of the gene was covered by the pB85-11recombinant clone. The presence of too many direct repeats in this region made the selection of sequencing primers extremely difficult. To overcome this situation the insert segment of this clone was further subcloned in smaller fragments to exploit the advantage of the universal primer sequences for the vector. For this purpose two primers, 85-E ($^{5'}$GTA TAC TTA CAG ATA GCA C$^{3'}$) and 50-E ($^{5'}$ GCC GAC AGT ATC ATT AAA C$^{3'}$), were used to amplify a 876 bp segment, using pB85-11 recombinant DNA as a template. The segment was cloned separately in a pCR™ II vector and designated as pCR85-11.1. The insert piece of pCR85-11.1 was restriction digested with Hind III enzyme and as a result of this, two DNA fragments of 4.3 Kb and 443 bp were produced. The 4.3 Kb fragment consisted of 495 bp insert piece and the rest of it (3.8 kb) was the plasmid vector part. This specific fragment was re-circularized to form the pCR85-11.1.1 subclone. The 441 bp fragment consisted of a 383 bp insert piece and a 60 bp plasmid piece. The 441 bp fragment was subcloned at the Hind III site of the pBluscript SK(-) phagemid and designated as pB85-11.1.2. The recombinant DNA of pCR85-11.1.1 was double digested with Hind III and EcoR I. The generated fragments were purified from the agar gel by the Gene clean technique and were further restriction digested with Sau3A I enzyme. The Sau3A I digestion generated two fragments of 317 bp and 247 bp. These fragments had a 9 bp and a 60 bp of plasmid sequence, respectively. These two pieces were separately subcloned at BamH I-EcoR I and BamH I -Hind III sites of pBluscript SK(-) phagemid. They were designated as pB85-11.1.1.1 and pB85-11.1.1.2 respectively.

Sequence of 50kD and 85kD recombinant clones:

Two vector primers from the opposite direction were used to reveal the complete

sequence of a 565bp insert fragment of the pB50-6.1 recombinant. Sequence analysis of this region did not indicate the presence of any possible reading frames for the 50kD antigen gene. The composite sequence analysis of the pB50-6.2 and pCR50-6.2.1 recombinants indicated a possible reading frame for the 50kD antigen gene present in the 2.2 kb fragment of pB50-6.2. The first methionine was located at the 848 bp downstream of the 5' end of this fragment, and the reading frame was continued all the way to it's 3' end. Further sequence analyse of the pB50-6.3.1 and pB50-6.3.2 recombinants revealed the complete sequence profile of the 50kD antigen gene.

Sequence analysis of the pB85-17 recombinant clone of the 90-12 strain helped to identify the presence of an 1155 bp uninterrupted reading frame of the 85kD antigen gene. However the fragment did not contain the 5' or 3' end of the gene. Further analysis of this clone revealed a partial sequence homology with the the 50kD gene of the 25D strain which helped in the selection of the two primers 50-C and 50-D for amplifing the 5' end of the gene. The sequence analysis of the cloned, amplified product (recombinant pCR85-3) revealed the 5' end of the gene. Analyse of the reading frames for subclones pCR85-11.1, pCR85-11.1.1, pCR-11.1.2, pB85-11.1.1.1 and pB85-11.1.1.2 exposed the complete 3' end sequence information of this gene.

Genomic localization of 50kD and 85kD strain specific antigen homologues:

The presence of a variable number of tandem repeats in the ORFs of both the 50kD and 85kD antigen genes, increases the possibility that these genes might be residing in a multigene family category. Thus there may be more than one copy of these genes, with other variable numbers of repeats, present somewhere in the chromosome. To confirm this, a specific probe of 1.5 kb was generated by PCR from the 90-12 genomic DNA following the procedure described above. The primers for amplification were selected in such a way that the amplified product contained a common 697 bp upstream and 180 bp downstream regions from the first methionine of the both 50kD and 85kD antigen genes. As a control, the insert segment of two other recombinant clones expressing the 55kD and 51kD antigen genes of the 90-12 strain were used as a probe.

The α^{32}P labeled probes were hybridized to the *E. risticii* genomic DNA of the 25D and 90-12 strains. The genomic DNA of both strains were digested with EcoR I and HinD III. Since Sau3A I was the single restriction enzyme used to obtain λ-ZAP recombinants, *E. risticii* DNA of both strains, digested with Sau3A I, were also used for the identification of homologous genomic DNA fragments in these recombinants.

The probes made with the inserts of the 55kD and 51kD recombinant clones of the 90-12 strain hybridized with the same- size fragments in each of the three restriction enzyme digests of both strains.

Molecular structure of *Ehrlichia risticii* SSA homologues:

The molecular structure of *E. risticii* of the SSA homologues (50kD and 85kD) antigen genes were revealed by analyzing the complete nucleotide and amino acid sequences of these two proteins. The complete nucleotide sequences of the genes were constructed from the sequences of individual clones and their subclones. Due to the presence of several direct repeats in these genes, the sequences obtained from the overlapping and adjoining clones and their subclones were further confirmed by amplification and sequencing of those areas directly from the genomic DNA of their respective strains.

Nucleotide sequence analysis:

A total of 2632 bp (25D strain) and 3357 bp (90-12 strain) were sequenced in the cloned *E. risticii* DNAs. The nucleotide sequence of the cloned 25D strain consisted of 869 bp of 5′ noncoding region , 1617 bp of the ORF, and 146 bp of a 3′ noncoding region. The nucleotide sequence of the cloned 90-12 strain consisted of 696 bp of a 5′ noncoding region, 2547 bp of the ORF, and 114 bp of a 3′ noncoding region. The base compositions of the sequenced DNAs showed high A + T contents (70%), especially in the 5' and 3' noncoding regions (71-80%). This reflects a high A + T-rich genomic DNA in *Ehrlichia*.

Structure of the 50kD antigen gene:

The nucleotide sequence of the 50 kD antigen gene ORF and 5′ and 3′ flanking

regions were determined and the amino acid sequence was deduced and depicted (Figure 3, SEQ ID NO: 3 and 4). An ATG translation start site at base pair position 175 and a TAA termination site at base pair position 1792 completed an ORF of 1617 nucleotides encoding 539 amino acids. The deduced sequence of the 50kD antigen has a calculated molecular mass of 59.829kD, which is in reasonably close agreement to the size originally observed on SDS-PAGE. The possible transcription initiation site and upstream control region are indicated in Figure 3. The upstream control region contained nearly perfect -10 and -35 consensus prokaryotic promoter sequences.

The ORF of the 50kD antigen gene continued uninterrupted at least 66 bp upstream of the proposed ATG translation start site. This 5′ region had no ATG codons present which could potentiate another translation initiation site. The further 5′ upstream region of this gene had two ATG codons which may be considered as translation initiation sites, but there were two distinct stop signal within 50 bases downstream of these two ATGs. Also, the recombinant in the expression vector produced a full-length product, while lacking the region 5′ of the proposed ATG. These two pieces of evidence nullified the possibility of these two ATGs as a translation initiator. The space between the -35 and -10 regions was 17 bp, which is consistent with the optimal spacing (17± 1) for prokaryotic promoters. The sequence GAAAAA at 7 bp upstream from the start codon was identified as a potential ribosome-binding site for m-RNA translation.

The non-coding region downstream of the translation termination site was a 143 bp stretch containing inverted repeats bordered by a thymine rich region, resembling prokaryotic *rho*-independent transcription terminators. These features are denoted in Figure 3.

Structure of the 85kD antigen gene:

The nucleotide sequence of the 85kD antigen gene ORF and 5′ and 3′ flanking regions were determined and the amino acid sequence was deduced (Figure 2, SEQ ID NO: 5 and 6). An ATG translation start site at base pair position 175 and a TAA termination site at base pair position 2722 completed an ORF of 2547 nucleotides encoding 849 amino acids. The deduced sequence of the 85kD antigen has a calculalated molecular mass of 94.333kD, which

is in reasonably close agreement to the size originally observed on SDS-PAGE. The possible transcription initiation site and upstream control region are indicated in Figure 2. The upstream control region , the translation start site, and first 178 bp after the first ATG were almost identical with the 50kD antigen gene sequence. The ORF of 85kD antigen gene continued uninterrupted at least 66 bp upstream of the proposed ATG translation start site. Like the 50kD antigen gene sequence, this 5′ region had no ATG codons present which could potentiate another translation initiation site, and it did not affect the full-length expression of the 85kD antigen, as the recombinant expression vector produced a full-length product while lacking the region 5′ of the proposed ATG. The proposed ribosome-binding site GAAAAA was present 7 bp upstream from the start codon.

The non-coding region downstream of the translation termination site was a 112 bp stretch containing inverted repeats bordered by a thymine rich region, resembling prokaryotic *rho*-independent transcription terminators. These features are denoted in Figure 2.

Repeat motifs and their nature in 50kD and 85kD SSA homologues:

The DNA sequence analyse of the 50kD and 85kD antigen genes revealed the presence of several direct repeats in both genes. The frequency of these repeats were more in middle of the genes and many of these repeats were identical in both genes. All these identical repeats coded for same amino acids but the position and the frequency of repetition were quite different in both genes.

Type of Repeats	Repeat Sequence	Repeated from Base
I	AAAGAAATACT	957, 1434, 777, 1287, 1353, 648.
II	GAAATACTCAC	807, 1356, 651, 1290, 1383.
III	AAATTTAAAGA	978, 1242, 852, 1110, 915.
IV	CTAAAAGAGAT	510, 1017, 891, 1149.
V	AAAGACATACT	501, 1071.
VI	TTTAAAGAGCT	342, 1113.
VII	ATTTTTTATAA	75, 119.
VII	AACTTTAAAGG	408, 1179.

IX	AAGTTTAAAGA	339, 1584.
X	TACTCACTAAT	457, 1504.
XI	AGTTTAAAAAA	669, 1309.
XII	ATAAGTTTAAA	237, 288

Table 1. Repeat locations and sequences along the 50kD antigen gene. Analyses were conducted on the 11-base repeats. There were 12 different types of 11-base repeats present in the complete sequence of the gene. A total of thirty-four 11-base direct repeats were identified in the gene.

Repeats in the 50kD gene:

There were a total of 97 repeats present in the 50kD antigen gene sequence. These repeats were not totally identical in their lengths and sequences. They were first categorized according to their lengths and then, under the same lengths they were grouped according to their sequence profiles. The minimum length of these repeats was 10 bases, whereas the maximum length was 38 bases. The result of these Analyses are represented in Table 1.

Repeats in the 85kD gene:

The structures of repeats in the 85kD antigen gene were almost identical to the 50kD antigen gene. There are a total of 356 repeats present in this gene sequence. As for the 50kD antigen gene, these repeats were categorized according to their lengths and then, under the same length, they were grouped according to their sequence profiles. The maximum and minimum lengths of these repeats were 55 and 10 bases respectively. As with the 50kD antigen gene the 11- mer repeats were also abundant in the complete sequence of the 85kD antigen gene, and they also were further analyzed for their specific positions in the sequence. The results of these analyses are presented in Table 2.

Type of Repeats	Repeat Sequence	Repeated From Base.
I	ATACTTACAGA	652, 1963, 1300, 901, 832, 385, 2260, 1729, 316.
II	AAATTTAAAGA	1984, 2116, 1852, 1390, 1252, 853, 784, 1915.
III	CTAAAAGAGAT	1891, 2023, 1429, 892, 760, 376, 1228.
IV	AAAGAAATACT	1696, 1567, 1165, 2227, 1030, 2161, 646.
V	TACTTACAGAT	1064, 1964, 834, 902, 1301, 1730, 2261.
VI	AAAGACATACT	310, 1945, 1351, 1282, 883, 814, 367.
VII	ACAGCTAAAGA	2275, 2302, 1771, 1744, 2302, 1159.
VIII	TTTAAAGAACT	1393, 2515, 2185, 856, 1323, 339.
IX	GAAATACTTAC	2164, 2257, 641, 1168, 1726.
X	AGCACTGGTAA	1975, 2005, 1381, 844, 1312.
XI	GATAAATTTAA	1912, 2380, 781, 1249, 1849.
XII	CTTATAGAAAG	934, 1333, 865, 349, 550.
XII	GAAATACTCAC	676, 2230, 1699, 1570, 1033.
XIV	ACCGGTAACTT	532, 916, 1433, 2047.
XV	ATGCAACAAAA	2204, 2621, 1007.
XVI	GCTAAAGAAGT	1189, 2278, 1747.
XVII	CTTACAGATAA	904, 2035, 1441.
XVIII	GCAATAACTGG	733, 1864.
XIX	ATGGTAAGGAC	494, 746.
XX	ACTTATAGAAG	417, 1401.

Table 2. Repeat locations and sequences along the 85kD antigen gene. Analyse were conducted on the 11- base repeats. There were 20 different types of 11 base repeats in the complete sequence of the gene. A total one hundred and one 11-base direct repeats were identified in the gene.

Analysis of deduced amino acid sequences of SSA homologues:

The amino acid sequence analyse of the 50kD and 85kD antigen genes indicated a considerable homology between these two SSA homologues. That the identical repeats of these two genes code for the same amino acids, indirectly indicates a conserved region between these two genes. From a comparison of the 32 amino acid sequences encoded in the N-terminal ends of the 50kD and 85kD antigens an almost identical signal sequence was identified for both proteins. Only one substitution of leucine for isoleucine occurred at residue 26 of the amino acid sequence in the 90-12 strain. These signal peptides for both strains consist of a polar region and a hydrophobic core, of which the same characteristics are seen in the signal peptides of other prokaryotic cells. The hydrophobic core region is extended from the 16th to 28th residues in the signal sequence. The predicted processing site of the signal peptide is at the bond between the 31st and 32nd amino acids, with isoleucine as the N -terminal amino acid of the mature SSA in both cases.

Amino acid sequence comparison of the SSAs of these two antigenic variants is presented in Figure 5 (SEQ ID NO: 4 and 6). In these analyses, substitution or the addition of one or several contiguous amino acid residues were identified throughout the molecules, but the significant homology in amino acid sequence of the 50kD and 85kD antigen was very pronounced in certain regions of the two molecules. These specific areas were designated as ID (identical domain) I-VIII in Figure 5. The most interesting feature of these IDs was the unique distribution of domains in the linear amino acid sequence of individual antigens. The domains were positioned one after another (ID I to ID VIII) in the 50kD antigen, whereas the positioning of the same domains was totally different in the 85kD antigen. In these ID regions, the similarities in the amino acid sequences between these two individual strains vary from more than 94% to less than 79%.

ID I is the largest identical domain, consisting of 129 amino acids. Here the amino acid sequence of the 50kD and 85kD antigens were very similar, and estimated homology is 89.15% (87.08% in nucleotide sequence) with 14 amino acid conversions. The position of this particular domain was the same in primary structures of both the antigens. This domain contained the signal sequence region of the the SSA homologues.

ID II consists of 51 amino acids. When comparing SSA homologues, this particular domain is found further downstream in the 85kD antigen. Here the estimated homology was 88.24% (89.54% in nucleotide sequence) with six amino acid conversions in between the 50kD and 85kD antigens.

ID III consists of 42 amino acids. The estimated homology is 92.85% (92.06% in nucleotide sequence) with 3 amino acid conversions. This particular domain is also found further downstream in the 85kD antigen as compared to the 50kD antigen.

ID IV consists of 21 amino acids. Here the estimated homology in amino acid sequence is 90.48% (85.71% in nucleotide sequence) with 2 amino acid conversions. With respect to the 50kD antigen this particular domain is found further upstream in the 85kD antigen.

ID V consisted of 39 amino acids. Among all the domains, this area had the minimum homology of 79.49% (80.34% in nucleotide sequence) in SSA homologues. In the 85kD antigen this domain is found further upstream as compared to the 50kD antigen.

The ID VI domain region has the maximum homology of 94.55% (93.82% in nucleotide sequence) between the two antigens. Similarly, the ID VII and ID VIII domains possess the high homology. ID VII has 92.11% homology (85.08% in nucleotide sequence) and ID VIII has 94.12% homology (96.73% in nucleotide sequence) in their respective areas of the SSA homologues.

After comparing the position of all the identical domains in SSA homologues it is clear that six domains out of eight are changed with respect to their positions in these antigens. In the 85 kD antigen the domains are further apart from each other as compared to the 50kD antigen, and these gaps are filled with new sequences. These observations indirectly indicate the generation of more new and different domains in the 85kD antigen.

Hydropathy analysis of SSA homologues:

Hydropathy analysis showed that the SSAs of both strains have alternative hydrophilic and hydrophobic motifs which are characteristic of transmembrane proteins. The hydropathy plot of the 50kD antigen revealed four major hydrophobic stretches which are

sufficient in length and hydrophobicity to serve as transmembrane domains. The largest hydrophobic stretch belongs to the identical domain I, and forms the hydrophobic core region of the predicted signal peptide. The other three hydrophobic stretches are clustered in last 60 amino acids of the C-terminus of the protein. Hydropathy analysis of the 85kD antigen indicats the presence of at least eight major hydrophobic regions. Any one of these regions can act as a transmembrane domain. Like the the 50kD antigen, this antigen also possesses the largest hydrophobic region in its identical domain I, and other three hydrophobic regions in the last 60 amino acids of the C-terminus. The other four major hydrophobic regions are distributed between residue200 and residue 410 in the sequence. Hydrophilicity indices for both antigens indicated the presence of many outer membrane domains which may be exposed on the outer surface of the organism or the inner cytoplasomic side of the membrane.

Epitope analysis of SSA homologues:

Locating the possible antigenic determinants by analyzing protein amino acid sequences in order to find the point of greatest local hydrophilicity, is a common technique nowadays. This was accomplished by assigning each amino acid a numerical value (hydrophilicity value) and then repetitively averaging those values along the peptide chain. The point of highest local average hydrophilicity was invariably located in, or immediately adjacent to, an antigenic determinant or epitope. Using this technique combined with analysis of the flexibility of proteins, the possible antigenic determinants of the 50kD and 85kD antigens were determined. Analysis of the comparative position of these epitopes in the common domains of the 50kD and 85kD antigens was critical to the evaluate the presence of possible cross-reactive and strain specific antigenic determinants in the 25D and 90-12 strains.

In order to compare the structural as well as antigenic aspects of the SSA homologues, Chou-Fasman predictions of the secondary structure of both the 50kD and 85kD complete antigens were plotted. None of these plots were identical to each other. Those regions predicted to have a high likelihood of antigenicity were also determined by the algorithm of Jameson and Wolf. Several regions of high antigenic indices appeared to be conserved in

both the antigens, although their positions and orientations in the secondary structure are quite different. Analysis of antigenicity of the 50kD indicated nine major areas with high antigenic indices (residues 76-80, 118-122, 274-278, 332-336, 362-366, 478-482, 508-512, 518-522, and 528-532). Among these nine major areas, the first two belong to ID-I; the 3rd one belonged to ID-IV; the 4th and 5th, to ID-VI; and the last four to an unique amino acid sequence region of the 50kD antigen which had no homology with the 85kD sequence. Analysis of antigenicity of the 85kD antigen indicated nine major areas with high antigenic indices (residues 76-80, 108-112, 118-122, 212-216, 246-250, 426-430, 590-595, 622-627, 844-848). Among those nine major areas, the first two belonged to ID-I, and the 3rd, 4th and 5th to ID-IV, ID-V and ID-VI respectively. The last three belonged to an unique amino acid sequence region of the 85kD antigen which had no homology with the 50kD antigen sequence. Several regions of high antigenic index in both antigens appeared to be conserved (residues 76-80, 118-122, 274-278, 332-336 in the 50kD antigen and 76-80, 108-112, 118-122, 212-216, 426-430 in the 85kD antigen). A high antigenic index region in the 85kD antigen belonged to ID-V, where as the ID-V in the 50kD antigen does not possess such type of high antigenic index region. This type of variation in this region of both the 50kD and 85kD antigens was predicted because the homology between the ID-V's in SSA homologues was minimum (79.49%) when compared to the other identical domains of these two antigens.

Recombinant antigens and their characteristics:

The complete ORF of the 50kD and 85kD antigens were constructed by PCR and cloned in pRSET-C expression vector. The correct ORF of the genes were confirmed by cloning and sequencing the PCR amplified product separately in pBluescript SK(-) phagemids prior to expression.

SK(-) recombinant clones of the 50kD and 85kD antigens:

The molecular size of the PCR generated fragments which contained the full length genes of the 50kD and 85kD antigens were 1.61kb and 2.54kb respectively. They were cloned separately in SK(-) phagemids. The BamH I - EcoR I restriction digestions of the

recombinant phagemids generated right size inserts, which were expected from the sequence information for these genes. Sequence analyse of these recombinant inserts confirmed the correct amplification of the SSA genes directly from their respective strains.

pRSET-C recombinant clones of the 50kD and 85kD antigens:

Total 18 positive pRSET-C recombinant clones of the 50kD and 85kD antigen genes (nine for each gene) were separately analyzed by restriction digestions to confirm the proper transfer of inserts from SK(-) phagemids to pRSET-C expression vectors. All nine positive clones from the 50kD recombinants were successfully transferred in expression vectors, whereas in the 85kD group only four of the clones were successfully recombined with the expression vectors. Finally, the complete 50kD and 85kD antigens were expressed in the pRSET-C systems. Coomassie Blue staining of expressed proteins indicated that maximum expression was achieved four to five hours after the IPTG induction.

Western blot analysis of the 50kD and 85kD expressed proteins:

The identities of the expressed proteins were established to be the 50kD and 85kD antigens by the reactivities of *E. risticii* (25D and 90-12 strains) polyclonal antisera and the 85kD clone specific antibody with the 50kD and 85kD antigens of their respective strains and corresponding expressed proteins. Both the 50kD and 85kD antigens migrated anomalously during electrophoresis and appeared to be 9.0kD smaller than the encoded sizes.

Example 2

Isolation of Strain Specific Surface Antigen Gene of
Ehrlichia risticii ATCC Type Strain

Using the procedures outlined in Example 1, the gene encoding the 50kDa SSA from the ATCC type strain was isolated. The gene sequence and the amino acid sequence encoded thereby is shown in Figure 4 (SEQ ID NO: 7 and 8).

Example 3

Challenge Experiments

SUMMARY

To study the role of major antigens of *E. risticii* in protective immune response, we expressed the genes of the 55 kDa, 51 kDa and 85 '50 kDa-strain-specific antigens of the 90-12 85 kDa antigen and 25-D (50 kDa antigen strains in *Escherichia coli.* Mice immunized with these purified recombinant proteins of *E. risticii* developed strong and specific humoral immune response. The recombinant 85 kDa antigen of the 90-12 strain protected mice against challenge infection with both *E. risticii* strains, whereas its homologue from the 25-D strain, the recombinant 50 kDa antigen, protected mice against only the homologous strain challenge, but not against the heterologous 90-12 strain. Sera from mice immunized with the 85- or 50-kDa antigens did not inhibit the replication of cell-free Ehriichia in in vitro neutralization assays. Sera from normal mice and mice immunized with other antigens caused non-specific neutralization of *E. risticii.* Immunoglobulin G from mice immunized with the 51 kDa protein of the 90-12 strain caused partial in vitro neutralization of both strains of *E. risticii.* These studies demonstrate that the 85/50-kDa-strain-specific antigen of *E. risticii* is involved in immunoprotection against PHF.

RESULTS

The protective capabilities of the purified recombinant antigens of *R. risticii* were tested in mice. In a pilot experiment, the 51 kDa, 55 kDa, 85 kDa, and 51 + 85 kDa antigens of the 90-12 strain were used to immunize the mice. Immunizations were performed by intraperitoneal inoculation of the respective antigen(s). The antibody response of mice to the recombinant antigens was determined by IFA using MM cells infected with the 90-12 strain. The prechallenge serum antibody titers of the different experimental groups are shown in Figure 6. The antibody titers varied from 1/40 to 1/640. The 85 kDa and 51 + 81 kDa groups of mice contained higher titers compared to the mice in the 51 kDa and 55 kDa groups. After the challenge infection with the 90-12 strain, the mice in 51, 85, 51 + 85 kDa, and the 90-12 organism groups did not show any clinical signs up to 21 days post-challenge. The 55 kDa

and adjuvant groups showed only mild clinical signs.

In a second experiment, the 50, 85, 51 + 85 kDa antigens of the 90-12 strain, and the 51, 50, 51 + 50 kDa antigens of the 25-D strain were included in the experimental groups. As positive controls, the mice were immunized with the purified organisms of the 90-12 and 25-D strains. The negative controls included the 55 kDa antigen, adjuvant, and uninoculated groups. At the time of the challenge infection, the serum antibody titers of these mice against the 90-12 strain (IFA titers) were obtained (Figure 7). After challenge infecting with the 90-12 strain, mice in the 85 kDa and 51 + 85 kDa groups showed significant protection (Figure 8). In the 85 kDa immunized group, only two out of eight mice suffered mild clinical signs for one and two days respectively. In the 51 + 85 kDa immunized group one mouse suffered from the infection and it died on day 12 post-challenge. The prechallenge serum antibody titer of this mouse was comparatively lower than the rest of the mice in that group. Mice in the positive control groups were completely protected from the infection. Mice in the negative control groups suffered from the clinical infection. The clinical signs of the mice immunized with either strain's 51 kDa antigen were less severe compared to those of the negative control groups.

In a third experiment, the mice were challenge infected with the 25-D strain. Even in the negative controls, the severity of the infection was less, thus confirming the lower pathogenicity of the 25-D strain. The 55 kDa immunized mice suffered mild clinical signs for only two days. None of the experimental groups showed any clinical signs.

DISCUSSION

The various challenge experiments described herein indicate that the recombinant strain-specific antigens, primarily the 85 kDa antigen of the 90-12 strain or the 90-12 strain itself can be used for immunization purposes. Any variants of *E. risticii* that bind to the antibodies to the 85 kDa antigen of the 90-12 strain may also be used for an attenuated bacterial vaccine. At present, vaccine effectiveness of existing PHF vaccines is low, and it is believed that the present invention can provide a superior vaccine against PHF.

Also, the antigens disclosed herein can be utilized in diagnostic tests or test kits to

diagnose PHF in horses. In addition, the nature of the repeated sequences of SSA can be used to generate intragenic primers to obtain specific DNA amplification finger printing (DAF) to differentiate various strains of *E. risticii*. The DNA amplification finger printing (DAF) of field *E. risticii* isolates are shown in Figure 9.

The following references are incorporated herein by reference in their entirety:

U.S. Provisional Application Serial No. 60/059,252, filed on September 18, 1997;

Biswas, Biswajit, Molecular basis of antigenic variation of strain-specific surface antigen gene of Ehriichia risticii and development of a multiplex PCR assay for differentiation of strains, Ph.D. Thesis, Univ. of Maryland, College Park, MD, USA SO (1996), 186 pp. Avail.: Univ. Microfilms Int., Order No. DA9707569 From: Diss. Abstr. Int., B 1997, 57(10), 6067;

Vemulapalli, Ramesh, Molecular analysis of differences between two strains of Ehriichia risticii and identification of protective antigen, Ph.D. Thesis, Univ. of Maryland, College Park, MD, USA, SO (1996) 176 pp., Avail.: Univ. Microfilms Int., Order No. DA9707676, From: Diss. Abstr. Int., B 1997, 57(10), 6125;

Vemulapalli et al, *Veterinary Parisitology*, 76, (1998), pp. 189-202; and

Dutta et al, *Journal of Clinical Microbiology*, Feb. 1998, pp. 506-512.

Obviously, numerous modifications and variations of the present invention are possible in light of the above teachings. It is therefore to be understood that within the scope of the appended claims, the invention may be practiced otherwise than as specifically described herein.